

EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1375
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-210-361-4

alignment_scores:
Quality: 140.00 Length: 606
Ratio: 0.507 Gaps: 32
Percent Similarity: 45.545 Percent Identity: 22.112

alignment_block:
US-09-303-518D-465 x US-09-210-361-4 ..

Align seg 1/1 to: US-09-210-361-4 from: 1 to: 1375

```
64 GCACACGCTCAGATTGGCAACAGATTCTTTATCCGGCAGGTTCTCGA 113
   :::::||||| :::::||||| :::::|||||
643 SerTyrAlaLeuLeuThrAsnLysSerSerValProArgValTyr 659
   :::::||||| :::::||||| :::::|||||
114 CGCTCAGCATTTGCACCCGACGCGGAATAC 144
   :::::||||| :::::||||| :::::|||||
659 rGlYAspMetPheThrAspAspLysGlnTyrMetAlaHisLysTrhIleA 676
   :::::||||| :::::||||| :::::|||||
145 .....CACGTATTC 153
   :::::||||| :::::||||| :::::|||||
676 snTyrGlnAlaIleGlnThrLeuLeuLysAlaArgIleLysTyrValSer 692
   :::::||||| :::::||||| :::::|||||
154 GCGACAGAGGGGAACTCCGACGCGCAGGTAT 189
   :::::||||| :::::||||| :::::|||||
693 GlYgLYglnAlaMetArgAsnGlnGlnValGlyAsnSerGlnIleIleTh 709
   :::::||||| :::::||||| :::::|||||
190 .....ATCGATTGGAAACATACAAAGCCATCAGTTGGCAACC 229
   :::::||||| :::::||||| :::::|||||
709 rSerValArgTyrGlyLysGlyAlaLeuLysAlaThrAspTrhGlyAspA 726
   :::::||||| :::::||||| :::::|||||
230 TGTTCAATCCAGAG.....GCGGCCATTAAAGAAATATCGCTAC 270
   :::::||||| :::::||||| :::::|||||
726 rGThrTrhArgThrSerGlyValAlaValIleGlnLysAsnProSer 742
   :::::||||| :::::||||| :::::|||||
271 ATT.....GTCCGCTTTCCGAT.....CA 290
   :::::||||| :::::||||| :::::|||||
743 LeuAlaGlnLeuLysAlaSerAspArgValValAlaAsnMetGlyAlaAlaHI 759
   :::::||||| :::::||||| :::::|||||
291 CGGAGCAGAGTGCATTCCTCCCTTC.....GACAAAC.....C 322
   :::::||||| :::::||||| :::::|||||
759 sLysAsnGlnAlaLysArgProLeuLeuThrTrhAspAsnGlyIleL 776
   :::::||||| :::::||||| :::::|||||
323 ATGCTCAGCATTCGATTCTGATGAAGCCGAGTCCCGTTGACGATTC 372
   :::::||||| :::::||||| :::::|||||
776 YAlaLysHisSerAspGlnGlnAlaIaGly..... 786
   :::::||||| :::::||||| :::::|||||
373 ACCCTTACCGCATTCATGGACGATACGAA.....CACATCCCGC 416
   :::::||||| :::::||||| :::::|||||
787 ...LeuValArgTyrTrhAsnAspArgLysGlnLeuIlePheThrAlaI 802
   :::::||||| :::::||||| :::::|||||
417 CGAC.....GGCTATGACGGCGCACAGCGCGCTATCC..... 453
   :::::||||| :::::||||| :::::|||||
```

```
802 aspIleLysGlyTyrAlaAsnProGlnValSerGlyTyrLeuGlyValT 819
   :::::||||| :::::||||| :::::|||||
454 ..GCTCCCAAGAGCGCGGAGGATATATACACTACGACATTAAGCGGTT 501
   :::::||||| :::::||||| :::::|||||
819 rPValProValGlyAla.....AlaAla 826
   :::::||||| :::::||||| :::::|||||
502 GCCCAAAATATCCGCTCAACCTGACCGCAACACCGCAGCAGCGGACAA 551
   :::::||||| :::::||||| :::::|||||
827 AspGlnAspValArgValAlaAlaSerThrAlaProSerTrhAspGlyL 843
   :::::||||| :::::||||| :::::|||||
552 GCTTGTCGACCGCTTTCACAATACCGGTAGTATGCTGACGCAAGAGTAG 601
   :::::||||| :::::||||| :::::|||||
843 sSerVal.....HisGlnAsnAlaAlaLeuAspSerArgValMetP 857
   :::::||||| :::::||||| :::::|||||
602 GCGACGATTCAAACGCGCCACCGATACACCCCGAGCTGCACAGATTC 651
   :::::||||| :::::||||| :::::|||||
857 heGlnGlyPheSerAsnPheGlnAlaPheAlaThrLysLysGlnGlyL 873
   :::::||||| :::::||||| :::::|||||
652 GGCAATGCCCGCGAAGCTTTCAC.....GGCAC 680
   :::::||||| :::::||||| :::::|||||
874 ThrAsnValValIleAlaLysAsnValAspLysPheAlaGlnTrpGlyA 890
   :::::||||| :::::||||| :::::|||||
681 TGCAGATATCTGTCAAAAACATATCGCGCGCGCAGCAAAATTTGTCGG 730
   :::::||||| :::::||||| :::::|||||
890 lThrAsp.....PheGlnMetAlaProGlnTyrValSer 902
   :::::||||| :::::||||| :::::|||||
731 CAGCGCATGCCGCTG.....CAGGGTATACGCAAGCTCAAAAT 771
   :::::||||| :::::||||| :::::|||||
902 erTrhAspGlySerPheLeuAspSerValIleGlnAsnGlyTyrAlaPhe 918
   :::::||||| :::::||||| :::::|||||
772 GCTGTATATGCAGCGCTGGGTCTGCTTTCACACGAAACAGATGGCGCG 821
   :::::||||| :::::||||| :::::|||||
919 ThrAspArgTyrAspLeuGly...IleSerLysProAsnLysTyrGlyTh 934
   :::::||||| :::::||||| :::::|||||
822 CATACAGATTTGGCAGAT...ATGGCGCACTCAAAAGACTATGGCGCAG 868
   :::::||||| :::::||||| :::::|||||
934 rAlaAspAspLeuValLysAlaIleLysAlaLeuHisSerLysGlyIle 951
   :::::||||| :::::||||| :::::|||||
869 CAGCAATCCGCGATTGGGACGTCGCAAAACCCCAATGCGCGCAAGGATA 918
   :::::||||| :::::||||| :::::|||||
951 yValMetAlaAspTrpValProAspGlnMetTyrAlaLeuProGlnLys 967
   :::::||||| :::::||||| :::::|||||
919 GAACCGCTCAGCAATATCTTACGCGAGTATCCCGCTCAAGGATGG 968
   :::::||||| :::::||||| :::::|||||
968 GluValAlaThr..... 971
   :::::||||| :::::||||| :::::|||||
969 AGCTGTTCG...GGAATATACGCTTGGCGGATCAGCGCACATCTCT 1015
   :::::||||| :::::||||| :::::|||||
972 .AlaThrArgValAspLysTyrGlyThr.....ProV 982
   :::::||||| :::::||||| :::::|||||
1016 TCAAGCGGTGCGAGATG...GGCGAGATCCGATTGCGGAAAGGAAATCG 1062
   :::::||||| :::::||||| :::::|||||
982 alAlaGlySerGlnIleLysAsnThrLeuTyrValValAspGlyLysSer 998
   :::::||||| :::::||||| :::::|||||
1063 GCCGTACGCGCAAT.....TTGCCGATGCGGC 1091
   :::::||||| :::::||||| :::::|||||
999 SerGlyLysAspGlnGlnAlaLysTyrGlyAlaPheLeuGlnGluLe 1015
   :::::||||| :::::||||| :::::|||||
1092 ATACGCAATACCCGCTTACCATTCGCCGAATATCGTTCAAACT 1141
   :::::||||| :::::||||| :::::|||||
1015 uGlnAlaLysTyrProGlnLeuPheAlaArgLysGlnIle..... 1028
   :::::||||| :::::||||| :::::|||||
1142 TGGAGCAGCTTACGGCAAAACATCACCTCTCAACCGTCCGCGCG 1191
   :::::||||| :::::||||| :::::|||||
1029 .....SerThrGlyValPromet 1034
   :::::||||| :::::||||| :::::|||||
1192 TCAACGGAAGATGTGAAA...CTGGCAACACAAACGCCACCGGAGAC 1238
   :::::||||| :::::||||| :::::|||||
1035 AspProSerValLysIleLysGlnTrpSerAlaLysTyrPheAsnGlyTh 1051
   :::::||||| :::::||||| :::::|||||
1239 CAAGATGCCCTTGAACGCTTAAGGTTTCCGATTTGAAAAAGACGTAA 1288
   :::::||||| :::::||||| :::::|||||
1051 rAsnIleLeuGlyArgGlyValaGlyTyr.....ValL 1062
   :::::||||| :::::||||| :::::|||||
```

```
1289 AATACGATACGAGATTAATTAACCGCTGTACACAACTGAATCTATAGAT 1338
      |||
1062 eulysaspglnalathrasnThr.....TyrPheSerLeuValSer 1075
      |||
1339 GAACCCGCTCTTAACTCTTAAGGTCTGTGCGATCGGTCAT.....TC 1382
      : : : : : ||| : : : : : |||
1076 AspAsnThrPheLeuProLysSerLeuValasProAsnHisGlyThrse 1092
      : : : : : ||| : : : : : |||
1383 TTGGCTATAATCTGCCAGATTCATACGCAAAATTATCCAGAGCAAGTA 1432
      - - - - - : : : : : |||
1092 rSerSerValThrGlyLeuValPheAspGlyLys.....GlyT 1105
      : : : : : |||
1433 GAATCGATATATATCCACCTAAATAATTACTCTCTTCACGACCGCTACCA 1482
      : : : : : ||| : : : : : |||
1105 yValTyrTyrSerThrSerGlyAsnGlnPheLysAsnAlaPheIleSer 1121
      : : : : : ||| : : : : : |||
1483 AAGAGCACTTAATATGATATTTGATTAATTTGTAATGATGACTTA 1532
      ||| ||| ||| : : : : : |||
1122 LeuGlyAsnAsnTrpTyrTyrPheAspAsnAsnGlyTyrMetValThr.. 1137
      : : : : : ||| : : : : : |||
1533 AGGTCATCAGACTAAAGCTCAAGATTTGATG.....GANG 1573
      ||| : : : : : ||| : : : : : |||
1138 .GlyAlaGlnSerIleAsnGlyAlaAsnTyrTyrPheLeuSerAsnGlyI 1154
      : : : : : ||| : : : : : |||
1574 TTCAATTGTCTAA.....ACAGAGAGAGCAACTTGGATG 1611
      : : : : : ||| : : : : : |||
1154 LeuGlnLeuArgAsnAlaIleTyrAspAsnGlyAsnValLeuSerTyr 1170
      : : : : : ||| : : : : : |||
1612 GCTAGTAGGATGTAG 1629
      : : : : : ||| : : : : : |||
1171 TyrGlyAsnAspGlyArg 1176

seq_name: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:US-09-199-637A-293

seq_documentation_block:
; Sequence 293, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 293
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-199-637A-293
```

```
alignment_scores:
  Quality: 134.00      Length: 280
  Ratio: 1.098         Gaps: 12
  Percent Similarity: 43.571  Percent Identity: 26.786
```

```
alignment_block:
US-09-303-518D-465 x US-09-199-637A-293
Align seg 1/1 to: US-09-199-637A-293 from: 1 to: 262
```

```
108 TCTGACCGCTCAGCATTTGCAACCCGACGGGAATATACCACTATTTCGCA 157
      : : : : : ||| : : : : : |||
27 AlaArgGlnAlaAlaAlaArgProGlyArg.....ValAlaGln 39
      : : : : : ||| : : : : : |||
158 GCAGGGGGGAACTTCCGAGCGCAGCGGTCATATTCGATTGG..... 200
      - - - - - ||| : : : : : |||
39 nGlyAlaGlyLysCysArgGlyGlnProAlaTrpArgTrpSerGlyLys 56
      : : : : : ||| : : : : : |||
201 .....AATCATCAAAAGCATATGTTGGGCAACCTGTCATACCA 239
      : : : : : ||| : : : : : |||
56 rProGlnArgArgGlyValAlaProGlyAlaGlyGln...ValHisArg 71
      : : : : : ||| : : : : : |||
240 GCAGCGCGCATTAAGAGAAATATCGCTACATTTCCGCTTTTCGATC 289
      : : : : : ||| : : : : : |||
72 ArgHisAspGlnAlaAlaArgGlyArgGlnAlaAlaArgProGlyAspArgse 88
      : : : : : ||| : : : : : |||
290 ACGGCGACGAGATCCATTCCTCCCTTCGCAACCATGCTTCATTCGAT 339
      ||| |||
88 rArgArgArg.....
      : : : : : ||| : : : : : |||
340 TCTGATGAAGCGGTATGCCGTTGACGATTCACGCTTACCGCATCA 389
      : : : : : ||| : : : : : |||
92 .....AspProProAspHisPro 97
      : : : : : ||| : : : : : |||
390 TTGGACGCGATACGAACACATCCCGCGACGGCTATGACGGGCCACAG 439
      : : : : : ||| : : : : : |||
98 GlyProAlaAlaAlaAspGlnGlnProGlyAlaAspArgArgThrAr 114
      : : : : : ||| : : : : : |||
440 GCGCGCGCTATCCGCTCCCAAGAGCGGAGGATATATACAGCTACAG 489
      ||| ||| ||| : : : : : |||
114 gArgArgGlnAspArgHisArgArgGlyProGly.....ProAlaHis 129
      : : : : : ||| : : : : : |||
490 ATAAAAGCGCTTCCCAAAATATCCGCT...CAACSTGACGACACACG 536
      : : : : : ||| : : : : : |||
129 rGlnArgArgSerAlaGlyArgProGlnGlnAlaProAlaValPro 145
      : : : : : ||| : : : : : |||
537 CAGCACCGGACACAGCGCTTTCGACCGTTTCCAAATACCGGTACTATGC 586
      |||
146 GlyHisGlyGly.....Al 150
      : : : : : ||| : : : : : |||
587 TGAACGACGAGTAGG.....CGACGATTA...ACGGGCCACCGA 627
      ||| ||| : : : : : ||| : : : : : |||
150 AspArgArgCysGlnValArgProArgArgValArgGlyThrProGlyLys 167
      : : : : : ||| : : : : : |||
628 TACAGCCCGGAGCTGACGATCGGGCAATGCGCGCAAGCTTTCACGG 677
      : : : : : ||| : : : : : |||
167 rProGlnArgThrGlnAlaGlyArgProGlyHisProValHisArg 183
      : : : : : ||| : : : : : |||
678 CACTGCAGATATCTCAAAACATCATCGCGCGACGAGAAATGTTCG 727
      : : : : : ||| : : : : : |||
184 ArgThrAlaHisHisGlyArgArgArgGlnGlyArgCysHisGlyArg 200
      : : : : : ||| : : : : : |||
728 GCGCAGCGCATGCGCTGCGAGGATTAAGCGAAG.....CTCAAC 768
      : : : : : ||| : : : : : |||
200 gArgGlnHisAlaGlnAlaGlySerGlyAlaArgArgAlaAlaLeuArg 217
      : : : : : ||| : : : : : |||
769 ATTGCTGTATGACGAGCTTGGCTGCTTCACCGAANA.....CAA 812
      : : : : : ||| : : : : : |||
217 rGlySerTyrTrpAlaArgValSerProValHisArgGlyLysArg 233
      : : : : : ||| : : : : : |||
813 GATGCGCGCATACAGATTTGGCAGATATGCGCACTCAAAAGCTATG 862
      ||| ||| ||| : : : : : |||
234 AlaGlyAlaProLeuProGlnGlyAlaGlyArgThrGlnArgGlyArg 250
      : : : : : ||| : : : : : |||
863 CCGCAGCAGCATCCCGCATTTGGCGCATCAAAACCCCA 902
      : : : : : ||| : : : : : |||
250 gHisHisArgHisPro.....ProTrpProGln 259
      : : : : : ||| : : : : : |||

seq_name: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:US-09-413-814-78

seq_documentation_block:
; Sequence 78, Application US/09413814
; Patent No. 6225064
```



```
1182 CGTCCCGCCGCTCAACGAGAAAGATGTGAACCTGGCAACCAACGCCACC 1231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
646 ArgSer:ProAlaGly.....ValAlaValAlaArgAlaGlyP 657
1232 CGAAGACCAACAGTCCGCTTGGACGCTAAAGGCTTCCGAATTTTGAAAA 1281
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
657 roGlyGlyArgAlaPro..... 662
1282 GAGGTAAATACGATACGAGAAATTAATACCGCTGTACACAGTGAATCC 1331
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
663 ...ValLeuAspAspGluAlaLeuGlyArgAlaAlaProGly..... 675
1332 TATAGATGAACCGCTTTAATCTTAAGGCTGTGTGCGATCGCTCAT 1381
675 ..... 675
1382 CTGGCTGTATTAAGTCCGAGATTCATACGCAAAATTAACCAAGCAAGT 1431
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
676 .....ProAlaGly 679
1432 AGAATCGATATATCCCACTAAATTAATCTCTTCAGACCGCTACC 1481
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
680 ArgProArgAlaGlyArgAlaAlaHisProArgAlaAlaGlyAlaGlyP 696
1482 AAAA.....GACCTAATTAATGATATTTGGATAAAT 1513
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
696 oArgArgAlaGlyArgAlaAlaHisProArgAlaAlaGlyAlaGlyP 713
1514 TTGGTAATGAATGAGCTAAAGTCCATCAGA 1545
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
713 LysLysAspLeuProLeuHisGlyProAlaArg 723

seq_name: /cgn2_6/protdata/1/aa/6B_COMB.pep:US-09-199-637A-351

seq_documentation_block:
; Sequence 351, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 351
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-199-637A-351

alignment_scores:
Quality: 130.00 Length: 473
Ratio: 0.756 Gaps: 21
Percent Similarity: 36.364 Percent Identity: 23.890

alignment_block:
US-09-303-518D-465 x US-09-199-637A-351
Align seg 1/1 to: US-09-199-637A-351 from: 1 to: 447
99 CCGGAGGTTCTCGACCGCTGACATTTTGAACCGGAGGAAATATCCACC 148

||||| ||||| ||| :|||
7 ProAlaThrSerArgLysProArgValAlaGlyProAlaValAlaArgPr 23
149 TATTCGGCAGCAGCGGGGAGACTTCCGAGCCGACGCGGTCAATATGGATG 198
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
23 oValProHisArgHisAlaProAlaGlyArgAlaPro..... 35
199 GGAACATATCAAAAGCCATCAGTTGGCAACCTGTTCATCCAGCA..... 242
||| ||| ||||| ||||| |||||
36 .....ProArgSerGly...ProValHisProAlaAlaPro 46
243 .....GGCGGCCATTAAAGAAATATCGCTACATTTGTCGCT 280
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
47 SerThrAspProGlyGlyHis.....AspHisArgValArgGlnProG 61
281 TTTCGGATCAGCGGCACGAGTCCATTCCTCCCTTGCAACACATCCCTCA 330
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 LysIAspArgProGlyAlaGlnGlyArgCysLeuArg..... 73
331 CATTCGGAATTCGATGAAGCCGCTAGTCCGTTGACGAGTTCAGGCTTTA 380
||| ||| |||
74 .....Leu 74
381 CCGCATTCATTTGGACGAGTACGAACACC.....AT 411
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
75 ProHisGlnThrGlyArgProArgSerLeuAlaGlyAlaGlyLysAsnArg 91
412 CCGCGCGACGCTATGACGCGGCCACAGGCGGCGGCTATCCCGCTCCCA 461
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 proThrLeuAlaGlnProGlySerArgGlySerLeuAlaGlyLysLysPro 108
462 AGCGCGGAGGATATATACAGCTACGACATAAAGCGCTGCCCAATA 511
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
108 roAlaArgArgVal.....AlaAlaAspAlaArg 117
512 TCCGCTCAACCTGACCGCAACCGCAGCAGCGCAACGCTGTGTGCAG 561
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
117 proAlaGlnProAspArgGlnAlaGlyAlaGlnProGlyAla.....G 132
562 CGTTTCCAAATACCGGTAGTATGCTGACGCGCAAGGATGAGGAGGAT 611
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 LysLeuHisGlnThr.....ArgValArgGlnArgGln 142
612 CAACGCGCCACCGCATACAGCCCGAGCTGSA..... 644
||| ||| ||||| ||||| |||||
143 GlyThrGlyGlyAlaProAspProArgAlaGlyAlaThrTyrArgAlaAl 159
645 ...CAGATCGGCAATGCCGCCGAAGCTTT..... 671
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
159 aValArgAlaGlyGlnLeuAlaArgAspSerLeuArgAlaAspGlyLysA 176
672 .....CAAGGCACTGCAGATAT..... 689
||| ||| ||| ||||| |||||
176 rGValLeuArgProGlnGlnArgGlnLeuHisTyrArgTyrArgGln 192
690 .....CGTCAAAACATCAT.....CGCGCGG 712
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
193 AlaGlyProValProGlyArgGlnArgTrpHisProValProAlaArgSe 209
713 CAGGAGAAATGTT..... 725
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
209 rArgArgProAlaAspGlyHisAlaGlyLysThrAlaProGlyAspProG 226
726 .....CGCGCAGCGAGTGCAGGAGTATACGGAAGGCTC 764
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
226 LysLysLysArgAlaArgGlyArgArgProAlaGlyGlyArgArgThr 242
765 AAACATTTCTGTATGACAGCGCTTGGCTGCTTCCACCGCAAAACAGA 814
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
243 CysAlaSerSerAlaProProThrArgThrSerProProLysSerAlaPr 259
815 TGGCGGCATCAACGATTTGGCAGATATGCGCACTCAAAACATATGCC 864
||||| ||| :|||
```

```

259 oglyalaserlalarthserthrtalaserthrsersercysa 276
865 GCAGCAGCCATCCGCGATTGGCAGTCCAAACCCCATGGCCGACAGAG 914
276 latyrthralalalargthrtproarglthrsproalalalargthrt 292
915 CATAGAGCCGTCAGCAATATCTTACGCGAGTCAATCCCGTCAAGGGA 964
293 Hisproglinalarproglarargthrsarproalaglaryglinalas 309
965 TTGAGCTGTGGGGAATA..... 986
309 pargalargthrtglglalagluculeuProleuproglyglinarpr 326
987 ..CGGCTGGGCGCATCCAGCGACA..... 1010
326 roarglagllyshslaglyalargleuProvalrargatarg 342
1011 .....TCCTGTCAAGCGGTCGACA 1030
343 Proaspralaserarprroalarglarygysalaglycsglnpr 359
1031 TGGCGAGATCGCATTCGCGAAAGGAATC..... 1061
359 oglyalargatrgluproglylthrsnarprglinarprglyleuprog 376
1062 ..CGCGTCAGCGCAATTTTGGCGATGCGCATACCGCAATACCGCTC 1109
376 llyrghlsaraglnalalaspalslaglythrtargllyasproleu 392
1110 .....CCCTTACCATTCGCGAAATATCGCTCAACT 1141
393 Gluprohisararglprroalarglproasprvalproleuasp.. 408
1142 TGGAGCAGCGTTACGCGAAAGAAA.....CATCAGCTCTCAACCGCG 1185
409 .....AlaleuproproglulysalaglyhisargluysVallysa 423
1186 CGCGCTCAACGAGAA 1202
423 rgrProvalrarglgn 428
seq_name: /cgn2_6/ptodata/1/iaa_5b.COMB.pep:US-09-041-886-23
seq_documentation_block:
: Sequence 23, Application US/09041886
: Patent No. 6235872
: GENERAL INFORMATION:
: APPLICANT: Bredesen, Dale E.
: APPLICANT: Rabizadeh, Sharoz
: TITLE OF INVENTION: Proapoptotic Peptides, Dependence
: TITLE OF INVENTION: Polypeptides and Methods of use
: NUMBER OF SEQUENCES: 72
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Campbell & Flores LLP
: STREET: 4370 La Jolla Village Drive, Suite 700
: CITY: San Diego
: STATE: California
: COUNTRY: United States
: ZIP: 92122
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/041,886
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Campbell, Cathryn A.
: REGISTRATION NUMBER: 31,815
: REFERENCE/DOCKET NUMBER: P-LJ 2626

```

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-9001
: TELEFAX: (619) 535-8949
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1185 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-041-886-23

```

```

alignment_scores:
  Quality: 130.00      Length: 685
  Ratio: 0.494        Gaps: 37
  Percent Similarity: 38.394      Percent Identity: 22.336

```

alignment_block:

US-09-303-518D-465 x US-09-041-886-23 ..

Align seg 1/1 to: US-09-041-886-23 from: 1 to: 1185

```

65 CACAGCGCTCAGATTGGCAACGATTTCTTATCCGCGAGTTCTGCAC 114
253 HisProPro.....ProthrtProIleSerValSerSerSergl 266
115 CGTACGATTTGCAACCGCAGCGGAATACCACTATTCGCGACGAGGG 164
266 yAlSerGlYalAProProthrlYsProProthrtThProValGlYgl 283
165 GGAACCTTGGCGAGCGCGGTCAATCGATTGGGAACATCAAGCC 214
283 lYasleuProSerAlAProProProAlasnphe.ProHisValThrPr 299
215 ATCAGTTGGGCAACCTGTATCCAGACAG..CGGCATTAAAGAAATAT 263
299 oAsleuProProPro.....ProAlaleuArgProleuAsnAlas 314
264 CGGCTACA.....TTGCGCGCTTTCCGATACAGCGGACGAGAGTCG 304
314 eAlaserProProglYleuGlYalaglInProleuProglY..... 327
305 ATTCCCGCTTCGACACCATGCTCACATTCGATTCGTGATGAAGCGGT 354
328 .....HisleuPro..... 330
355 AGTCCCGTTAGGATTCACGCTTTACCGCATTCATTTGGACGAGATACGA 404
331 .....SerProtyr.....AlameGlYInglYmetG 340
405 ACACCATCCGCGCGAGGTATGACGGGCCACAGAGGGGGGCTATCCG 454
340 lYglYleuProProglYProglulYsclYProthrlleuAlAProSerPro 356
455 .....CTCCCAAGCGCGAGGATATATACAGCTACGACATATAAGGC 498
357 HisSerleuProProAlaserSerSerAlAProAlAProPro...Metar 372
499 GTTCCCAAAATATCCGCTCAACCTGACCGACACGACGACGAGACA 548
372 gPneProtyrSerSerSerSerSerSerSerAlAAlAAlaserSerSer 389
549 ACGGCTTGCAGCGTTCCCAATACCGGTATATGTCGACCCAGAGAG 598
389 eSerSerSerSerSerSerAlaserProhe..... 399
599 TAGCGACGAGATTCAACGCGCCACCCATACAGCC...CGAGACTGAGC 645
400 .....ProAlaserGlAlAleuProSerlyrPr 409
646 AGATCGGCGCATCCGCGAGAGCTTTCAACGCGACGTGACGATATGCTCA 695
409 oHisSerProProPro.....ThrsleuSerValSerAsnG 423

```

```

696 AAMCATCATGCGCGCGGAGAGAAATGTGGCGGAGCGATCCCGTGC 745
    ::::: ||| ::::: ||| ::::: |||
423 InProProLyProProLyProProLyProSerLeuProSerLeuVal 437
746 AGGATATAGCGAAGGCTCAACATGTGTATGACAGCGCTGGGCTTG 795
438 .....TTPserGI 440
796 CTTTCACGCAAAAAA.....AGATGGCGCGCATCAAGCATTT 833
    ||||| ::::: ||| ::::: |||
440 nglYProProLyProProLyProProLyProLyLeuAlaAsnSer.... 455
834 GCGAGATATGGCGCAATCAAGACATGATGCGGAGAGCGATCCGGAT 883
456 .....AsnAlaHisProGIyProProLyProProSerLeuGI 467
884 GCGAGATCAAAACCCCAATGCGG.....CACAGGCAATAGAA 921
    ::::: ||| ::::: ||| ::::: |||
468 AlaGIserThrAlaHisProProLyAlaSerThrHisHisHisHisGI 484
922 GCGGTACGAAATCTTTACGCGATCATCCCGTCAAGAGATGGAGC 971
    ::::: ||| ::::: ||| ::::: |||
484 InGIInGIInGIInGIInGIInGIInGIInGIInGIInGIInGIInGI 500
972 TGTTCGGGAAATACGCGCTTGGG..... 995
501 AsnSerGIyProProLyAlaProProHisProLeuGIInGI 517
996 .....CGGCATCAGCGCATCT..... 1014
517 ySerSerHisHisAlaHisProLyAlaMetSerProSerLeuGIySer 533
1014 ..... 1014
534 LeuAlaProLyProProGIyProAlaHisLeuProProHisSerGI 550
1015 .GTCAACGCGTCCGAGATGGCGAGATCGCATGGCCGAAGGAAATCCG 1063
    ::::: ||||| ::::: ||| ::::: |||
550 nValSerTySerGIInAlaGIyProAsnGIyProProValSerSers 567
1064 CC...GTACGCGACAAATTTGGCGATGGGCGATAC..... 1095
    ::::: ||||| ::::: |||
567 eRAsnSerSerSerSerThrSerGIInGIySerTyProCySerHisPro 583
1096 .....GCCAATACCGCTCCCTTAACCTTC 1121
584 SerProSerGIInGIyProGIInGIyAlaProLyProProHePro..ProValP 600
1122 CCGAATAA.....TCCGTTCAACTTGGAGCAGCGTTAGCGCA..... 1159
    ::::: ||| ::::: ||| ::::: |||
600 roPIInValIThrThrSerSerAlaThrIleuSerThrValIleAlaThrVal 616
1160 .....AGAAAACATCACCTCTCTCAACCGCTCCG 1188
617 AlaSerSerProAlaGIyTyTyThrAlaSerProProGIyPro..Pro 632
1189 CCGTCAACGGAAGAAATGTGAACCTGGCAACAAACGCCACCGC..... 1233
    ||| ::::: ||| ::::: |||
633 Pro.....TyGIyLysAlaProSerPr 641
1234 .....AGACCAAGTCCGTTTGAAGGTAAGGTT..... 1266
    ||||| ::::: ||| ::::: |||
641 oGIyAlaTyTyThrAlaThrPro.....ProGIyTyLysProG 655
1267 .....CCGAATTTGAAAAAAGCAATTAATCGATAGCAGA..... 1302
655 LySerProProSerProHeArGIyThrProProGIyTyAlaGIyThr 671
1303 .....ATTAAATCCGCTGACCAAGCAAGTAA 1328
672 SerProProAlaGIyProGIyThrPheLySerProGIySerProThrValGI 688

```

```

1329 TCCATATAGTAGACCCGCTTTAAATCTAAAGT..... 1362
    ||| ::::: ||| ::::: |||
688 yProGIyProLeuProProProAlaGIyProSerGIyLeuProSerLeuPro 705
1363 .....TCTGCGATGGGCTCATCTTGGTGTATTAAGTCCACA 1401
    ::::: ||| ::::: |||
705 roProProAlaAlaProAlaSerGIyProProLeuSerAlaThr..... 719
1402 ATTCAATACGCAAAATTTACCAAGGCAAGTAGAATCAATATATCCACC 1451
    ::::: ||| ::::: |||
720 .....GlnIleLysGIInGIyProAl 726
1452 TAAAAATTAAC...TCTCCTCAGCAGCGCTACCA.....AAAGACCTA 1492
    ::::: ||| ::::: ||| ::::: |||
726 agIInGIyTyGIyThrProGIySerProValProProAlaArgSerPro 743
1493 AT..... 1494
743 erProProProLyAlaValAspValProSerHisAlaSerGIInSerAla 759
1495 .....AATGATATTTGGATTAATTTGTAAT..... 1521
760 ArgPheAsnLysHisLeuAspArgGIyPheAsnSerCysAlaArgSers 776
1521 ..... 1521
776 pleuTyThrPheValProLeuGIInGIySerTyLeuAlaLysLysAlaVal 793
1521 ..... 1521
793 spleuValGIInLysValArgArgGIInAlaGIInArgAlaArgGIIn 809
1522 .....GAATGACTAAGAGTCCATCAAGAACTAA 1550
810 LysGIInArgGIInArgGIInArgGIInArgGIInLysGIInArgGIInArgGIIn 826
1551 AGGTCAAGAAATTTGAATGGGATGTTCAATGTCTAAACAGCAGACA 1596
    ::::: ||||| ::::: ||| ::::: |||
826 sGIInArgGIInLeuGIInArgSerValLysLeuAlaGIInGIInGIIn 841

```

seq_name: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:US-09-351-200-2

seq_documentation_block:

```

: Sequence 2, Application US/09351200
: Patent No. 6320033
: GENERAL INFORMATION:
: APPLICANT: BOURBONNATS, Yves
: APPLICANT: LAMARRE, Claude
: TITLE OF INVENTION: CANDIDA ALBICANS GENE (CSA1) ENCODING A
: TITLE OF INVENTION: MYCELIAL SURFACE ANTIGEN, AND USES THEREOF
: FILE REFERENCE: 6013-71"US" CC/
: CURRENT APPLICATION NUMBER: US/09/351,200
: CURRENT FILING DATE: 1998-07-09
: EARLIER APPLICATION NUMBER: CA2,237,134
: EARLIER FILING DATE: 1998-07-10
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 1203
: TYPE: PRT
: ORGANISM: Candida albicans
: FEATURE:
: NAME/KEY: PEPTIDE
: LOCATION: (0)...(0)
US-09-351-200-2

```

alignment_scores:

```

Quality: 126.00 Length: 447
Ratio: 0.558 Gaps: 16
Percent Similarity: 50.559 Percent Identity: 19.687

```

alignment_block:
US-09-303-518D-465 x US-09-351-200-2 ..

Align seg 1/1 to: US-09-351-200-2 from: 1 to: 1203

```

32  TGTGCTACTGCGAGTGTGCTCCGATGCATGCACAGCCCTCAGATTGG 81
   |||||||
782  CysProThrPaspThrGlyCysLeuCysIleMetPro..... 794
82  GCAACGATCTTTATCCGCGAGGTTTCGACCGTCACGATTTGCAACC 131
   |||||||
795  .....ThrPheIleGlyAlaIleGlySerCysIleAlaGlyLeu 808
132  CGACGGGAATATACCATTTTCGCGACGAGGGGGAATTCGCGACGCA 181
   |||||||
808  ysIsglyGlnGluValIleSerValThrSerLeuGlySerIleCys 824
182  GCGGTATATCGATTGGCAACATCAAAAGCCATCATGTTGGCAACCTG 231
   |||||||
825  SerValAlaGlyValThrPaspProTyr.....TrpMetLeuPr 837
232  TTCATCCAGCGCGCGCATTTAAGGAATATGCGCTACATTTGCGCTT 281
   |||||||
837  oAlaAsnValGlnSerSerLeu...AsnAlaAlaIleThrAlaValAla 853
282  TTCGATCAGCGGCGACGAGTCCATTCCTCCGACAAACCATGCTCAG 331
   |||||||
853  hrSerAspSerAlaSerGluValAlaSerAlaSerGluSerAlaSerGln 869
332  ATTCCCATTTCTGATGAGCCGCGTACGCTTGACGATTCAGCCTTAC 381
   |||||||
870  ValPro.....GlnGluThrSerAl 876
382  CGCATCATTTGGGACGATACGACACATCCCGCGCGAGGCTATGAGCG 431
   |||||||
876  aAlaSerSerGlnSerAlaAsnSerValAlaSerAlaAlaIleProSer 893
432  GCCACAGGCGGCGGCTATCCGCTCCCAAGCGCGAGGATATATACA 481
   |||||||
893  snSerSerValSerAlaIleProSerSerAsnSerSerGlyAlaProAla 909
482  GGTACGACATTAAGGCG.....TGCCCAAAATATCCGC 516
   |||||||
910  AlaIleProSerAsnSerSerGlyAlaSerValIleProSerGlnSerAl 926
517  CTCACACTGACCGACACGACGACGCGACGCTTTCGACCGCTT. 565
   |||||||
926  aAsnAsnSerSerAlaSerAlaIleProSerAsnAsnSerSerAlaI 943
566  .....TCCCAATACCGGTA 580
943  leSerGlySerValAlaIleProSerSerTyrGlyAsnSerThrIleAla... 958
581  GTATGCTGACGAGGAGTAGGCGGATTCACAGCGCGCCACCGCATAC 630
   |||||||
959  .....GlnProSerThrSerThrIleThrSerAspAlaAlaSerIleTh 972
631  AGCCCCGAGCTGACAGATCGGCGAATCCCGCGGAG..... 667
   |||||||
972  rGlyProIleThrThrAspIleValIleThrAsnGluSerGlyIleValP 989
667  ..... 667
989  heThrSerThrValIleIleThrHisValSerGluTyrCysAspGlnThr 1005
   |||||||
668  CTTTCACGCGCGACTGCAGATATCTCA.....AAACATATCGCGCG 711
   |||||||
1006  SerAlaAlaIleValGlnSerSerAlaCysGluGlnGlnSerSerAla 1022
   |||||||
712  GCAGAGAAATTTGTGCGCGAGCGGATCCCGTCAGGGGTATTAAGGAG 761
   |||||||
1022  serGlnGlnAlaSerAlaSerSerGlnGlnValIleVal..... 1035

```

```

762  CTCAAACATGCTGTATGACAGGCTGTGCTGCTTTCCACGAAACA 811
   |||||||
1036  .....IleThrSerValValIleP...CysGluSerSerIleGln 1048
812  AGATGCGCGCGATCAACGATTTGGCAGATATGCGCAACTCAAGACTAT 861
   |||||||
1048  erIleGluSerValIleThrSerIleGluAlaIleHisIleThrGluVal 1064
862  GCCGAGACGCGATTCGCGATTGGGACGTCCAAACCCCAATGCCGACA 911
   |||||||
1065  IleAlaSerCysAlaSerGluLeuSerSerLeuSerSerAlaIleSerG 1081
912  AGGCATAGACCGCTCAGCATATCTTT.....ACGGCATCA 949
   |||||||
1081  uAlaMetIleThrValIleSerSerIleValGluValGlnIleSerAla 1098
950  TCCCGTCAAAGGATTTGAGCTGTTCGCGGAAATACGCTTGGCGCGC 999
   |||||||
1098  lAlysGlnThrSerLeuAlaIleValGlnSerSerAlaIleSerValGln 1114
1000  ATCAGC...GCACATCCCTGTCACGCGTCCGACATGGCGGATCGATT 1046
   |||||||
1115  leuSerAlaAlaHisAlaGlnIleSerSerSerGluAlaIleValAla 1130
1047  GCCGAAAGGAAATCCGCGCTCAGCGACACAAATTTGCCGATGGCGATAC 1096
   |||||||
1131  .....GlnThrAlaVal.....AlaGluAlaSerIleYsa 1140
1097  CCAAAATACCCGCTCCCTTACCATTCGCAAAATATCCGTCAAATCGGAG 1146
   |||||||
1140  laGlyAspGluIleSerThrGluIleValAlaSerIleThrIleThrVal 1156
1147  CAGCGTTACGCGCAAGAAACATCACCCTCCTCAACCGTCCGCGCTCAA 1196
   |||||||
1157  ser.....GlyIleGluThrGlyValSerIleGlnAlaIleValAla 1171
1197  CGGAAAGATGTGAACCTGCGCAAAACCAACGCCACCGGAG 1236
   |||||||
1171  nThrHisSerValAlaIleAlaAsnMetAlaAsnThrIle 1184

seq_name: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:US-08-783-774-2

seq_documentation_block:
; Sequence 2, Application US/08783774
; Patent No. 6054130
; GENERAL INFORMATION:
; APPLICANT: Spaete, Richard
; TITLE OF INVENTION: Jackman, Winthrop
; TITLE OF INVENTION: NON-SPLICING VARIANTS OF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/783,774
; FILING DATE: 15-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7682-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090

```


TELEFAX: 212-869-8864
 TELETYPE: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 907 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-783-774-2

alignment_scores:
 Quality: 125.00 Length: 566
 Ratio: 0.504 Gaps: 31
 Percent Similarity: 43.816 Percent Identity: 23.322

alignment_block:
 US-09-303-518D-465 x US-08-783-774-2 ..

Align seg 1/1 to: US-08-783-774-2 from: 1 to: 907

```

95  TTAATCGGCGAGTTCTGACCGTACGACATTGCAACCCGAGGAAATAC 144
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
379  ILeSerGIyAlaPheAla.....SerAsnArGThrPheAsp11 391
145  CACCTATTGGCGAGCGGGGAACTTCCGAGCGAGCGGTCATATCCG 194
   |||  |||  :::::  |||  :::::  :::::  :::::  :::::
391  eThrValSerGIyLeu...GIyThrAlaProLySThrLeuIleIleThra 407
   |||  |||  :::::  |||  :::::  :::::  :::::  :::::
195  ATTGGGAACATACAAAGCCATAGTTGGGCAACCTGTTTCATCCAGCAG 244
   |||  |||  :::::  |||  :::::  :::::  :::::  :::::
407  rGThrAlaThrAsnAlaThrThrThrIleSylsValIlePheSerIys 423
245  CGGCCATTAAAGAAATATCGCTACATTCCGCTTTCCGATCAGCGG 294
   |||  |||  :::::  |||  :::::  :::::  :::::  :::::
424  AlaPro.....GIuSerThrThrThrSerProThrLeuAsnThrThrGI 438
295  C.....ACGAAGTCCATCCCGCTTCGACAAACCATGC 326
438  yPheAlaAspProAsnThrThrThrGIyLeuProSerSerThr..... 452
327  CTCACATTCGATTCGATGATGAGCCGCTAGTCCGCTTGACGATTCAGCC 376
   |||  |||  :::::  |||  :::::  :::::  :::::  :::::
453  ..HisValPro.....ThraSLeuThr 459
377  TTTACCCATTCATTTGGAGCGATAGACAAACATCCCGCGAGGCTAT 426
   |||  |||  :::::  |||  :::::  :::::  :::::  :::::
460  AlaProAlaSerThrGIyProThrValSerThrAlaAspValThrSerPr 476
427  GAGC.....GGCCACAGGGGGGCTATCCGCTCCCAAGAGCGGAG 470
   |||  |||  :::::  |||  :::::  :::::  :::::  :::::
476  oThrProAlaGIyThrThrSerGIyAlaSerProValThrProSerProS 493
471  GGATATATACAGCTACGACATATAAGGCGTTGCCCAAAATATCCGCTCA 520
   :::::  :::::  |||  :::::  |||  :::::  :::::  :::::
493  erProThrAspAsnGIyThrGIuSerIysAlaProAspMetThrSerSer 509
521  ACCTGACCGACAAACCGGACAGCGGAGCGCTTGACGCTTTTCAC 570
   |||  |||  :::::  |||  :::::  :::::  :::::  :::::
510  ThrSerProValThrThrProThrProAsn.....AlaThrSerPr 533
571  AATACCGGTATGATGTCGACGAAGAGATGAGCGGATTCAAACCGCG 620
   |||  |||  :::::  |||  :::::  :::::  :::::  :::::
523  oThrProAlaValThrThrProThrProAsnAlaThr...SerProThrP 539
621  CACCCGATACAGCCCGAGCTGACAGATCGGGCAATGCCGCG..... 664
   |||  |||  :::::  |||  :::::  :::::  :::::  :::::
539  roAlaValThrThrPro.....ThrProAsnAlaThrSerProThrLeu 553
665  ..AAGCTTTACAGCGACTGACATATCGTCAAAAACATCATCGGCGG 711
   |||  |||  :::::  |||  :::::  :::::  :::::  :::::
554  GIySThrSerProThrSerAlaValThrThrProThrProAsnAlaThr 570

```

```

712  GCAGGAGAAATTTGGCGGCGAGGATCCGTCGAGGATTAACGAAG 761
   :::::  |||  :::::  :::::  |||  :::::
570  rSerProThrLeuGIyLysThrSerProThrSerAlaValThrThrProt 587
762  CTCAAACATTCGCTTATGACAGCGCTGGGCTGCTTCCCA..... 802
   :::::  |||  :::::  |||  :::::
587  hrProAsnAlaThrSerProThrLeuGIyLysThrSerProThrSerAla 603
803  .....CGAAACAGAGTGGCGGCATCAACGATTTGGCAGATATG 843
   |||  |||  :::::  |||  :::::
604  ValThrThrProThrProAsnAlaThrGIyProThrVal..... 616
844  GCGCAACTCAAAAGACTATCCGACAGCAGCATCCGCGATT..... 883
   |||  |||  :::::  |||  :::::
617  ....GlyGIuThrSerProGIuAlaAsnAlaThrAsnIleThrLeug 631
884  ..GGCAGTCCAAACCCCAATG.....CCGACACAGGACATAG 919
   |||  |||  :::::  |||  :::::
631  LyGIyThrSerProThrProValValThrSerGIuProLyAsnAla... 646
920  AAGCGTCAGCAATATCTTAAGCGAGTCATCC..... 952
647  ..ThrSerAlaValThrThrGIyGIuIleSAsnIleThrSerSerSerTh 662
953  .....CCGTCAAGGAGATTGGAGCTTCCGCGGANA 983
662  rSerSerMetSerLeuArGIuProSerSerAsnProGIuThrLeuSerProS 679
984  ATAGCGCTTGGCGGCGATCAGCGCAC.....ATCCTGACAGCGGTCCG 1027
   |||  |||  :::::  |||  :::::
679  erThrSerAspAsnSerThrSerHisMetProLeuIleThrSerAlaHis 695
1028  AGATGGCGAGATGCGCATTCGCGAAGGAAATCCGCGCTCA...GCCAG 1074
696  ProThrGIyGIyGIuAsnIleThrGIuValThrProAlaSerIleSerTh 712
1075  AATTTGGCGATGCGGATACGCGCAATACCGGCCCTTCCATTCGCG 1124
   |||  |||  :::::  |||  :::::
712  rHisIleValSerThrThrSerSerProGIuProArGIyThrThrSerG 729
1125  AATATTCGCT.....TCAACTTGGAGACGCTTACGCGCAAGAA 1164
   :::::  |||  :::::  |||  :::::
729  IuAlaSerGIyProGIuAsnSerSerThrSerThrLyGIuProGIyAl 745
1165  AACATACCTCCTCAACGCTGCCGCGTCAACGGAAGAAATGTGAACCT 1214
   |||  |||  :::::  |||  :::::
746  AsnValThrLyGIyThr...ProProGIuAsnAlaThrSerProGIuAl 761
1215  GGCAAACAAACGCCACCGAGAGACCAAAAGTCCG.....TTTGAGGTA 1258
   :::::  |||  :::::  |||  :::::
761  aProSerGIyGIu.....LySThrAlaValProThrValThrSerThrG 776
1259  AAGGCTTCCGAAATTTGAAAAAGACGTAAATAGCATACG..... 1299
   |||  |||  :::::  |||  :::::
776  LyGIyLySAlaAsnSerThrThrGIyGIyLySHisThrThrLySIsly 792
1300  ...AGATTAATACCGCTGTACCACAGATGATCTATAGATGAAACCGCT 1346
   |||  |||  :::::  |||  :::::
793  AlaArGIyThrSerThr.....GIuProTh 800
1347  CTTTAATCTTAAAGTTTCTGTGCGATCGGCTCATCTTGCTATATATG 1396
   :::::  |||  :::::  |||  :::::
800  rThrAspTyrGIyGIyAsp.....SerThrThrP 810
1397  CCAGAATTCAATACGCAAAATATACCAAGCAAGTAGAATCATGATATC 1446
   |||  |||  :::::  |||  :::::
810  roArGIy.....ProArGIyAsnAlaThrThrTyIleu 820
1447  CCACCTTAAATATTAATCTTCCTTCAGAGACGCTACCAAAAGACGTAA 1496
   |||  |||  :::::  |||  :::::
821  ProProSerThrSerSerLySLeuArGIuPro..... 830

```

```

1497 TGGATATTGGATAATTTGGTAATGATGACT...AAAGTCACATCAA 1543
831 .....:|||||
1544 GAACCTAAGGTCAAGATTGATGGATGCTCAATGCTCTAAACAGCA 1593
839 althrhAlaInAla.....ThValProValProProthSer 852
1594 AGAGACCA.....CTTGATGGGCTGTGT 1617
853 GlhProArhPheSerAsnLeuSerMetLeuValLeuGlnTrpAlaSer 866

seq.name: /cgn2_6/ptodata/1/1aa/PC/TUS-04611A-19
seq.documentat_block:
: Sequence 19, Application PC/TUS9504611A
: GENERAL INFORMATION:
: APPLICANT: Spaele, Richard and Jackman, Winthrop, T
: TITLE OF INVENTION: Non Splicing Variants of gp350/220
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
: STREET: 5 Palo Alto Square
: CITY: Palo Alto
: STATE: California
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/04611A
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/229,291
: FILING DATE: April 18, 1994
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Luann Caser
: REGISTRATION NUMBER: 31,822
: REFERENCE/DOCKET NUMBER: AVIR-003/000S
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-843-5163
: TELEFAX: 415-857-0663
: TELEX: 380816 COOLEYPA
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 907 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US95-04611A-19

alignment_scores:
Quality: 125.00 Length: 566
Ratio: 0.504 Gaps: 31
Percent Similarity: 43.816 Percent Identity: 23.322

alignment_block:
US-09-303-518d-465 x PCT-US95-04611A-19 ..
Align seg 1/1 to: PCT-US95-04611A-19 from: 1 to: 907

95 TTAATCGGAGGTCTCGACGCTGACGATTGAGACCGGAGGGAATAC 144
:::|||||
379 IIEsetGlyAlaIheAla.....SerAsnArhTrhPheAspI1 391
145 CACCTATTGCGAGCGAGGAGGACTTCCGACGCGAGGTCATATGCG 194
||| ||||| ||||| |||||

```

```

391 ethrValserGlyLeu...GlyThrAlaProLysThrLeuIleIleIhe 407
195 ATGGGAACATPACAAAGCCATCAGTTGGGCAACCTGTCATCCAGAG 244
407 rghrAlaIheAsnAlaIheTrhTrhTrhTrhTrhTrhTrhTrhTrh 423
245 CGGCATTAAGAAATATCGGCTACATTCGCCCTTCCGATCAGG 294
424 AlaPro.....GluSerTrhTrhTrhTrhTrhTrhTrhTrhTrh 438
295 C.....ACGAAGTCATTCCTCCCTTCGACAAACATATGC 326
438 yPheAlaIheProAsnTrhTrhTrhTrhTrhTrhTrhTrhTrhTrh 452
327 CTCACATTCGATTCGTATGAGCCGATGTCCTCCGATTCAGCC 376
453 ..HisValPro.....ThrAsnLeuTrh 459
377 TTATCCGATCCATTTGGAGAGATACGACACATCCCGCGAGGCTAT 426
460 AlaProAlaSerTrhGlyProTrhValSerTrhAlaIheValTrhSer 476
427 GAGC.....GGCCACAGGCGCGGCTATCCGCTCCCAAGGCGGAG 470
476 oThrProAlaGlyTrhTrhTrhSerGlyAlaSerProValTrhProSer 493
471 GGATATATACAGCTACGACATAAAGCGTTGCCCAAAATATCCGCTCA 520
493 erProTrhPaspAsnGlyTrhGluSerLysAlaProAspMetTrhSer 509
521 ACCTGACGACACACGCGAGCAGCAGCGGCTGTGCGACGCTTCCAC 570
510 ThSerProValTrhTrhTrhProTrhProAsn.....AlaTrhSer 523
571 AATACCGGTAGTATGCTGACGCAAGAGTAGCGGAGATTCAAAGCGC 620
523 oThrProAlaValTrhTrhProTrhProAsnAlaTrh...SerProTrh 539
621 CACCGATACAGCCCGAGCTGAGACATCGGCAATGCCCG..... 664
539 roAlaValTrhTrhPro.....ThProAsnAlaTrhSerProTrhLeu 553
665 ...AAGCTTTCACGCGCATGCAAGATATGTCAAAACATCATGCGCGC 711
554 GlyTrhSerProTrhSerAlaValTrhTrhProTrhProAsnAlaTrh 570
712 GCAGGAAATATGCGGCGAGCGAGCGATCGCGTACAGGTTAAGCGAAG 761
570 rSerProTrhLeuGlyLysTrhSerProTrhSerAlaValTrhTrhPro 587
762 CTCAAACATTCGTGTATGCACGCGTTGGCTGCTTCCA..... 802
587 hrProAsnAlaTrhSerProTrhLeuGlyLysTrhSerProTrhSerAla 603
803 .....CCGAAACAGATGCGCGCATCAACGATTTGGCAGATATG 843
604 ValTrhTrhProTrhProAsnAlaTrhGlyProTrhVal..... 616
844 GCGCACTCAAGACTATGCCGAGCAGCAGCCATCCGGAT..... 883
617 .....GlyGluTrhSerProGlnAlaAsnAlaTrhAsnTrhLeuG 631
884 ..GGGAGTCCAAACCCCAATG.....CCGACACAGGCAATAG 919
631 LysGlyTrhSerProTrhProValTrhSerGlnProLysAsnAla... 646
920 AAGCCGTCACATATCTTACGCGATCC..... 952
647 ...ThSerAlaValTrhTrhGlyGlnHisAsnIleTrhSerSerSerTrh 662
953 .....CCGTCAAAGGATGAGAGCTGTCGGGGA 983
662 rSerSerMetSerLeuArhProSerSerAsnProGluTrhLeuSerPro 679

```

```

984 ATACGGCTTGCGCGCATCAGCGAC.....ATCCTGTCAGCGGCTCC 1027
   ::::::::::::::::::::|||
679 erthseraspanserthserthserthserthserthserthserth 695
   ::::::::::::::::::::|||
1028 AGATGGCGGATCGATGCGCGAAGGAAATCCCGCTCA...GGAC 1074
   ::::::::::::::::::::|||
696 Prothrglyglylualsnietherthserthserthserthserth 712
   ::::::::::::::::::::|||
1075 AATTTGGCGATCGCGCATACCGCAATACCGCTTACCATTCCG 1124
   ::::::::::::::::::::|||
712 rthshsvalserthserthserthserthserthserthserth 729
   ::::::::::::::::::::|||
1125 AATATCCGT.....TCAACTGGAGCGAGTTCAGCGCAAGAA 1164
   ::::::::::::::::::::|||
729 lna1aserthserthserthserthserthserthserthserth 745
   ::::::::::::::::::::|||
1165 AACATCAGCTCTCAACCGTCCGCGTCAACGAAAGATGTAAGT 1214
   ::::::::::::::::::::|||
746 AsnValthrlysglythr...Prothglina1athserthserth 761
   ::::::::::::::::::::|||
1215 GGCAGCAACGCCACCGCGAAGCAAGTCCG.....TTGACGGTA 1258
   ::::::::::::::::::::|||
761 aproserglyln...LysThrAlaValProthValthrserthc 776
   ::::::::::::::::::::|||
1259 AAGGTTCCGAATTTGAAAAGAGTAAATACGATACG..... 1299
   ::::::::::::::::::::|||
776 lylglylala1asnerththththththththththththth 792
   ::::::::::::::::::::|||
1300 ...AGAATTAATACCGCTGACCAAGTGAATCTATAGATGACCC 1346
   ::::::::::::::::::::|||
793 Alaarghserthr.....GluProth 800
   ::::::::::::::::::::|||
1347 CTTAATCTTAAGCTCTGCGATCGCTCATCTTGTCTATATCTG 1396
   ::::::::::::::::::::|||
800 rthasptlyglylyasp.....Serththrp 810
   ::::::::::::::::::::|||
1397 CCAGATTCATACGCAAAATTCACAGCGAGTAGAATCAGATATAT 1446
   ::::::::::::::::::::|||
810 roArg.....Prothrglyrsna1aththththththth 820
   ::::::::::::::::::::|||
1447 CCACCTAAATATCTCTCTCAGCAGCGCTACCAAAAGACCTATA 1496
   ::::::::::::::::::::|||
821 Prothserthserthserthserthserthserthserth 830
   ::::::::::::::::::::|||
1497 TGGATATTGGATTAATTTGGTAATGATGACT...AAGGTCATCA 1543
   ::::::::::::::::::::|||
831 .....ArgtrpThrPheThrSerProth 839
   ::::::::::::::::::::|||
1544 GAATTAAGTCAGATTTGAATGGATGATTCATATGCTAAACAGGA 1593
   ::::::::::::::::::::|||
839 althtrthala1a1a.....ThyValProvalProththser 852
   ::::::::::::::::::::|||
1594 AGAGAGCAA.....CTTGATGGGCTACT 1617
   ::::::::::::::::::::|||
853 GluProthrgheserthserthserthserthserthserthser 868
   ::::::::::::::::::::|||

```

seq_name: /cgn2.6/ptodata/1/1aa/5B_COMB.pep:US-08-471-119A-2

```

seq_documentation_block:
; Sequence 2, Application US/08471119A
; Patent No. 5827706
; GENERAL INFORMATION:
; APPLICANT: Lettner, Ernst
; APPLICANT: Schneider, Elisabeth
; APPLICANT: Schoergerdorfer, Kurt
; APPLICANT: Weber, Gerhard
; TITLE OF INVENTION: Cyclosporin Synthetase
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5827706artis Corporation
; STREET: 59 Route 10
; CITY: East Hanover

```

```

STATE: New Jersey
COUNTRY: USA
ZIP: 07936
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,119A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kassenoif, Melvyn
REGISTRATION NUMBER: 26,389
REFERENCE/DOCKET NUMBER: 100-8029/CONT/CONT
TELEPHONE: 201 503 8474
TELEFAX: 201 503 8807
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 15281 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Tolypocladium niveum
STRAIN: ATCC 34921
US-08-471-119A-2

```

alignment scores:

Quality	Ratio	Length	Gaps
123.50	0.454	544	27
Percent Similarity: 50.000	Percent Identity: 19.853		

alignment block:

US-09-303-518d-465 x US-08-471-119A-2 ..

Align seg 1/1 to: US-08-471-119A-2 from: 1 to: 15281

```

112 GACCGTCACATTTGCAACCC...GACGGAAATACCACTATTGGCAG 158
   ||||| ::::::::::|||
4940 AsparGValargValargProlysasphneGlnIleGluPhePheGlyAr 4956
   ::::::::::|||
159 CAGGGGGAACCTTCCGAGCGCGCTCAT...ATCGATTTGGAAACA 205
   ::::::::::|||
4956 gmetAspGlnGlnIleIleIleArgGlyHisArgIleGluProAlaGluV 4973
   ::::::::::|||
206 TACAAAGCATCAGTTGGGCAACCTGTTTCATCCAGAGCGGCCTTTAA 255
   ::::::::::|||
4973 alGluGlnAlaPheLeuasnaspGlyPheValGluAspValAlaIle... 4988
   ::::::::::|||
256 GGAATATCGGCTACATGTCGCGCTTTCCGATCAAGGCGACAGAA.... 300
   ::::::::::|||
4989 .....ValIleArgThrProGluAsnGlnGluProGluMe 5000
   ::::::::::|||
301 .....GTCCATTCCTCCCTTGACAAACCATGCTTCATTCGATT 340
   ||| ::::::::::|||
5000 tValAlaPheValThrAlaIleGlyAspAsnserAlaIa... 5014
   ::::::::::|||
341 CTGATTAACCGCGGTACTCCCGTTGACGATTTAGCTTACCGCATCAT 390
   ::::::::::|||
5014 lueGluGlnAlaThrThrGlnIleGluGly.....TrrpAlaHis 5027
   ::::::::::|||
391 TGGGAGGATACGAAAC...CATCCGCGCA 419
   ::::::::::|||
5028 PheGluGlyAlaIleAlaAsnIleGluGluIleGluSerGluAlaIle 5044
   ::::::::::|||
420 CGGCTATGACGGGCCACAGGGCGGCGGTATCCGCTCCCAAGCGCGCA 469

```

```

|||||
5044 uGlyTyrAspMet.....GlyTrpThrSerMetTyrAspGlyT 5058
470 GGGATATATACGCTAGACATAAAGCGCTGCCCAATATCCGCTC 519
5058 hrGluIleAspLysAspGluMetArgGluTrpLeuAsnAspTrpMetArg 5074
520 AACCTGACCGCAACACCGACACCGGCAACGGCTTGTCAGCGTTTCCA 509
5075 SerLeuLeuAspGlyLysProAlaGly..ArgValLeuGluValGlyTh 5090
570 CAATACCGGAGT...ATGCTGACGCAAGAGATAGCGGACGATTCAAC 616
5090 rGlyThrGlyMetLeuMetPheAsnLeuGlyArgSerGlnGlyLeuGlu 5107
617 CGCGCCACCGCATACGCCCGGACGCTGACAGATCG.....GGCAAT 657
5107 rGlyTrpIleGlyLeuGluProAlaProSerAlaAlaGluPheValAsn 5123
658 GCCGCGCAACCTTTCACGCGCTGACATATGTCACAAAACATCATCG 707
5124 AlAlaLysSerPheProGlyLeuAlaGlyArgAlaGluValHisValG 5140
708 CGCGGCGAGGAATTTGCGCCGACGCGATGCCGTCGACGAGTAAAGC 757
5140 YThrAlaAlaAspVal.....GlyThrLeuGlnGlyLeuThr. 5152
758 AAGGCTCAACATTTGCTTATGACGCGCTTGGT...CTGCTTCCAC 804
5153 ....SerAspMetAlaValIleAsnSerValAlaGlnTrpPheProThr 5167
805 GAAACACAGATGGCGCGC..ATCAACGATTTGGCAGATATGGCGCACT 851
5168 ProGluTrpLeuAlaGluTrpIleLysSerLeuValGlnValProGly 5184
852 CAAGACTATGCGCGACGACCATCGCGATTTGGCGATGCCAACCCCA 901
5184 LysAlaGlnIleTyrLeuGlyAspMetLysSerTrpAlaMetAsn..... 5198
902 ATGCCGCAACAGGATAGAGCCGTCAGCAATATCTTTACGCGACGTATC 951
5198 ..... 5198
952 CCCGTAAGAGGATTTGAGCTGTCGGGAAATACGCTGGCGCGCAT 1001
5199 ....ArgAspPheAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 5213
1002 CACGCGACATCCTGTCAGCGGTCGACAGATGGCGCATTCGCTCCGA 1051
5213 nAlaSerLysAspArgValArgGlnLysMetGluLeuGluLysG 5230
1052 AAGGAAATCCCGCTGACGACCAATTTTGGCGATCGGCAATACGCCAA 1101
5230 LngLngLnuLeuLeuValAspProAlaPhePheThrAla..... 5242
1102 TACCCCTCCCTTACCATCCGAAATATCCGTTCAACTTTGGAGCAGCG 1151
5243 .....LeuAlaSerGlnLeuGlnAspArg 5250
1152 TTACGGCAAGAAACATCAGCTCTCAACCGGCGCGCGCTCAAGGAA 1201
5250 gIleGlnHisValGlnIleLeuProLysArgMetLysAlaTrpAsn.... 5265
1202 AGAATGTAAACTGGCAACAAACGCCAC..... 1230
5266 .....GluLeuSerSerTyrArgTyrAlaAlaValLeuHisLieser 5279
1231 .....CCGAAGACCAAACTGCCGTTTGACGTTAAAGGTTTCCCAATT 1274
5280 AspGluProLeuProIleTyrLysLysLeuAspProGlnAlaTrpIleAsn 5296
1275 TGAAGAAAGACGTAAATACGATACGGAAT...AATACCGCTGTACAC 1321
|||||

```

```

5296 eGlu.....GlySerArgLeuThrArgGluAlaLeuAlaG 5308
1322 AAGTAATCTTATAGATGAACCGCTTTATCTTAAGGTTCTGTCGA 1371
5308 LnuValLeuLysGluAsnGlu..... 5314
1372 TCGGCTCATCTTGTGCTATATACCTCCAGAAATTCATACGCAAAATTCC 1421
5315 AsnAlaGluSerValAlaIle...SerAsnIleProTyrSerLysThrVa 5330
1422 AAGCAAGGTAGATCATATATATCCACCTAAATTAATCTCTCTCAG 1471
5330 LnuAlaGlu.....ArgHisLeuValArgSerLeuAspGlnLysAs 5344
1472 CACCGCTACCAAGACCTAATATGATATTTGGATTAATTTGGTAT 1521
5344 LAsnAlaProGlnLnuSerMetAsp.....GlySer 5354
1522 GAATGACTAAAGGTCATCAAGACTAAGGTCAAGAAATTGAA..... 1566
5355 AspTrpLieserAla...ValArgThrArgAlaGlnGlnCysHisThrLe 5370
1567 .....TGGATGTTCAATTTGCTTAAACAGAAAGAGAGC 1600
5370 uSerAlaSerAspLeuPheAspIleAlaGluAspAlaGlyPheArgValG 5387
1601 AACTGATGGGCTGTAGGATGTATGAT 1632
5387 LnuValSerTrpAlaArgGlnHisSerGlnHis 5397

```

seq_name: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:PCT-US96-03916-6

seq_documentation_block:

Sequence 6, Application PCT/US9603916

GENERAL INFORMATION:

APPLICANT: Wild, Martha A.

APPLICANT: Cochran, Mark D.

TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS

TITLE OF INVENTION: AND USES THEREOF

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/03916

FILING DATE: 23-MAR-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/126,597

FILING DATE: 24-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 39116-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 985 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US96-03916-6

alignment_scores:

Quality: 122.50 Length: 670
 Ratio: 0.447 Gaps: 34
 Percent Similarity: 40.896 Percent Identity: 22.090

alignment_block:
 US-09-303-518D-465 x PCT-US96-03916-6 ..

Align seg 1/1 to: PCT-US96-03916-6 from: 1 to: 985

```

55  CCGATGCGATGACACAGCGCTCAGATTGGCAACGATCTTTATCCGGCA 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
322  ProTyrArgAlaLeuGlySerAsnValProArgAspSerIleArgpr 338
105  GGTTTCGACCGTCAGCATTTGAA.....CCGACGGGAATATCC 145
338  oGlyAlaThrLeuProPheAspThrAlaAlaProasp.PheAspThr 354
146  ACCATTCGGCAGCAGGGGGAACTGCCGAGCGCAGCGTC..... 187
355  GlyThrSerProThrProThrValProGluProAlaIleThrThrIe 371
188  .....ATATCGGAT.....TGGAAACATACAAAGCA 215
371  uIleProArgSerThrSerAspMetGlyPhePheSerThrAlaArgAlar 388
216  TCAGTTGGGCAACCGTTTCATTCAGCAGCGCGCATTAAGAAATATCG 265
388  hrGlySerGluThrLeuSerVal.....ProValGlnGluThrAsp 401
266  GCTACATTCGCGGTTTCGATCAG..... 292
402  ArgThrLeuSerThrThrProLeuThrLeuProLeuThrProGlyIuse 418
293  .....GGCAGCAAGTCCATTCGCC 311
418  rGluAsnThrLeuPheProThrThrAlaProGlyIleSerThrGluThrP 435
312  CTTCG.....ACAACATGCCCTCACATTCGATTCGATG 346
435  roSerAlaAlaHisGluThrThrGlnThrGlnSerAlaGluThrVal 451
347  AAGCGGTAGTCCGCTGACGGATTCAGC..... 376
452  PheThrGlnSerProSerThrGluSerGluThrAlaArgSerGlnSerG 488
377  .....TTACCGCATCCATTGGGAGCGATACGACACCATCC 413
468  nGluProThrPyrThrThrGlnThrProSerThrGluGlnAlaLeuT 485
414  CGCCGAGCGGTATGAGCGCCACAGGGCGGCTATCCGCTCCCAAG 463
485  hrGlnThrGlnIleAlaGluThrGluAlaLeuPheThrGlnThrProSer 501
464  GCGCGA.....GGCATATATACA..... 481
502  AlaGluGlnMetThrPheThrGlnThrProGlyAlaGluThrGluAlaPr 518
482  .GCTAGCAGATAAAGCGTGGCCAAATA.....TCCGCCCTCAA 521
518  oAlaGlnThrProSerThrIleProGluIlePheThrGlnSerArgSerT 535
522  CCTAGCGACACAC.....GCAGACCGGACACAGCGCTTG 556
535  hrProProGluThrAlaArgAlaProSerAlaAlaProGluValaPheThr 551
557  TCGACCGTTTCCACAATACCGGTAGTATGCGACGCAAGAGTAGGAGCAG 606
552  GlnSerSerThrValThrGluValaPheThrGln.....Th 564
607  GGATTCAAACGGCCACCGCATACAGCCCGAGCTGGACAGATCGGGCAA 656

```

```

564  rProSerThrValProLysThrThrLeuSerSerSerThrGluProAlaI 581
657  TGGCGCCGAGCGTTCAACGGCA...CTGCAGATATGCTCAAAAACATCA 703
581  IePheThrArgThrGlnSerAlaGlyThrGluAlaPheThrGlnThrSer 597
704  TCGCGCGCGCAGAGAAATGTGCGCGCAGCGATCCGCTGAGGGTATA 753
598  SerAlaGluProAspThrMetArgThrGlnSerThrGluThrHisPheP 614
754  AGCGAAGGCTCAACATTGCTGTATGCAAGCGCTTGGTCTGCTTCCAC 803
804  CCAAAACAAAGATGGCGCGCATTCACAGATTGGCAGATATGGCGCACTCA 853
622  rOlyAlaThrGlnThrProSerThrGluProGluValLeuThrGlnSer 638
854  AAGACT..... 859
639  ProSerThrGluProValProPheThrArgThrLeuGlyAlaGluProGl 655
860  .ATGCCGACAGCCATCCGCGATTGGCGAG.....TCCAAA 896
655  uIleThrGlnThrProSerAlaAlaProGluValaThrArgSerSers 672
897  CCCCAATGCCGCAAGGATAGAAAGCCGTACAGCATATCTTTACGGCAG 946
672  erThrMetProGluThrAla.....Gln 679
947  TCATCCCCGTCAAGAGGATTGAGCTGTCCGGGAAATACGGCTTGGCC 996
680  SerThrPro.....Le 683
997  GGCATCACGGCAGATCCTGTCAAGCGTGCAGATGGCGAGATCGCAT 1046
683  uAlaSerGlnAsnProThrSerSerGlyThrGlyThrHisAsnThrGluP 700
1047  GCCGAAGGGAATCCGCGCTCAGCGACATTTCCGATGGCGGATACG 1096
700  roArg.....ThrTyProValGlnThrThr 708
1097  CCAAAATACC.....CGTCCCTTACCATTCGCCGA 1125
709  ProHisThrGlnLysLeuTyThrGluAsnLysThrLeuSerPheProTh 725
1126  AATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAGAAACATACCTC 1175
725  rValValSerGluPheHisGluMetSerThrAlaGluSerGlnThrPro. 741
1176  CTCACACCGTCCCGCGCTCAACGGAAGAAATGTGAACGTGGCAACAAAC 1225
742  Leu.....LeuAspValLysIleValGlu..ValLysPheSerAsn... 754
1226  GCCACCCGAAGACCAAAAGTCCGTTTGAAGGATTAAGGATTTT 1275
755  .....AspGlyGluValThrAlaThrCys 762
1276  GAAAAAGCGTAAA.....TACGATACGAGAAATTAATACCGCTGACC 1319
763  ValSerThrValLysSerProTyrArgValGluThrHisThrLysValAs 779
1320  ACAAGTGAATCCATATGATGAAACCGCTTTAAT...CCTAAAGGTTG 1366
779  pleuValAspValMetAspGluIleSerGlyAsnSerProAlaGlyValP 796
1367  TCGATCGGCTCATCTCTTG.....TCTATACTGCAGAAATCAATAC 1410
796  heAsnSerAsnGluLysTrpGlnLysGlnLeuTyTyThrArgValThrAsp 812
1411  GCAAAATTCCAAGCA..... 1428

```

```

813 GlyArgThrSerValGlnLeuMetCysLeuSerCysThrSerHisSerPr 829
1429 .....GGTAGATCGATATA 1444
829 OGluProTyrCysLeuPheAspThrSerLeuIleAlaArgGlnIlyAspI 846
1445 TCCACCTAAATAATTACTCTCTTCACGACCG..... 1476
846 leAlaProGlnLeuTyrPheThrSerAspProGlnThrAlaTyrCysThr 862
1477 .....CTACCAAAAGGACCTAATAATGATATTGGATATAATTGGTAA 1520
863 IleThrLeuPro.....SerGlyValValProArgPhe..... 873
1521 TGATG.....ACTA 1531
874 .GluTrpSerLeuAsnAsnValSerLeuProGluTyrLeuThrAlaThr 890
1532 AAGTCCATTAAGACTAAAGTCAAGATTTGAATGGATGTCATTTG 1581
890 hValValSerHisThrAlaGlnSerThrValTrp..... 902
1582 TCTAAACAGAGAGACAACTTGATGGCTAGTAGGATGTAGCA 1631
903 LysSerSerAlaArgAlaGlnIlyAlaTrpIleSerGlyArgGlnGlyAs 919
1632 TTGA 1635
919 nile 920

```

seq_name: /cgn2.6/plodata/1/1aa/PCTUS.COMB.pep:PCT-US96-03916-66

```

seq_documentation_block:
: Sequence 66, Application PC/TUS9603916
: GENERAL INFORMATION:
: APPLICANT: Wild, Martha A.
: TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
: NUMBER OF SEQUENCES: 72
: CORRESPONDENCE ADDRESS:
: ADDRESS: Cooper & Dunham LLP
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/03916
: FILING DATE: 23-MAR-1995
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/126,597
: FILING DATE: 24-SEP-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 39116-A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 278-0400
: TELEFAX: (212) 391-0525
: INFORMATION FOR SEQ ID NO: 66:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 985 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
PCT-US96-03916-66

```

alignment_scores:

Quality: 122.50 Length: 670
Ratio: 0.447 Gaps: 34
Percent Similarity: 40.896 Percent Identity: 22.090

alignment_block:

US-09-303-518D-465 x PCT-US96-03916-66 ..

Align seq 1/1 to: PCT-US96-03916-66 from: 1 to: 985

```

55 CGGATGATGACAGACGCGCTGATTTGGCAACGATCTTTATCCGGCA 104
||| ||||| ||||| ||||| |||||
322 ProTyrArgAlaLeuGlySerAsnValProArgAspSerIleArgPr 338
105 GGTTCGACCGCTCAGCATTTGGAA.....CCGACGGGAAATACC 145
338 OGlyAlaThrLeuProProPheAspThrAlaAlaProAsp.PheAspThr 354
146 ACCGATTCGGGAGCGAGGGGGAAGCTTGGCAGCGAGCGGTC..... 187
355 GlyThrSerProThrProThrThrValProGluProAlaIleThrIle 371
188 .....ATATCGAT.....TGGGAAACATACAAGCA 215
371 uIleProArgSerThrSerAspMetGlyPhePheSerThrAlaArgAla 388
216 TCAGTTGGGCAACCTGTTGATCCAGACGCGGCATTAAGGAATATCG 265
388 hGlySerGluThrLeuSerVal.....ProValGlnGluThrAsp 401
266 GCTACATGTCGCGCTTTCCGATCAG..... 292
402 ArgThrLeuSerThrThrProLeuThrLeuProLeuThrProGluGlu 418
293 .....GGCAGCAAGTCCATTTCCC 311
418 gLysAsnThrLeuPheProThrThrAlaProGlyIleSerThrGluThr 435
312 CTTCG.....ACAACATGCTCATTTCCGATTCGATG 346
435 roSerAlaAlaIleGluThrThrGlnThrGlnSerAlaGluThrValVal 451
347 AAGCGGTAGTCCGCTTGACGATTCAGCC..... 376
452 PheThrGlnSerProSerThrGluSerGluThrAlaArgSerGlnSer 468
377 .....TTTACCGCATCCATTTGGAGCGATACGACACCATCC 413
468 nGluProTyrTrpPheThrGlnThrProSerThrGluGlnAlaLeu 485
414 CGCGGACGCTATAGCGGCGCACAGGGCGGTATCCGCTCCCAAG 463
485 hGlnThrGlnIleAlaGluThrGlnAlaLeuPheThrGlnThrProSer 501
464 GCGCGA.....GGCATATATACA..... 481
502 AlaGlnGlnMetThrPheThrGlnThrProGlyAlaIleThrGlnAlaPr 518
482 .GCTACGACATAAAGCGTTGCCCAAAATTA.....TCGCGCTCAA 521
||| ||| ||||| ||||| |||||
518 oAlaGlnThrProSerThrIleProGluIlePheThrGlnSerArgSer 535
522 CCGACCGGACGAC.....GCAGACCGGCAACGCGCTTG 556
535 hProProGluThrAlaArgAlaProSerAlaAlaProGluValPheThr 551
557 TCGACCGTTCCACATACCGGTAGTCTGACGCAAGAGTAGAGGAC 606
||||| ||||| ||||| |||||
552 GlnSerSerThrValThrGlnValPheThrGln.....Th 564
607 GGATTCAAAGCGCGCCAGCTACAGCCCGACCTGGACAGATCGGCA 656

```

564 rProSerThrValProLysThrThrLeuSerSerSerThrGluProAla1 581
567 TGGCGCGGAGGCTTCAACGGCA...CTGCACATATCGCAAAACATCA 703
581 LePheThrIgtThrGlnSerAlaGlyThrGlnAlaPheThrGlnThrSer 597
704 TGGCGCGGAGGAGAAATTTGCGGCGAGGATGCGCGTGCAGGGATA 753
598 SerAlaGluProAspThrMetArgThrGlnSerThrGlnThrHisPhe 614
754 AGCGAAGGCTCAACATTCGTGTATGCAGCGCTGGGTCTGCTTCCAC 803
614 eThrGlnAlaProSerThrVal.....P 622
804 CGAAACAGATGCGCGCATCAAGATTTGGCAGATATGGCGCAATCA 853
622 rLysAlaThrGlnThrProSerThrGluProGluValLeuThrGlnSer 638
854 AAGACT..... 859
639 ProSerThrGluProValProPheThrArgThrLeuGlyAlaGluProG1 655
860 .ATGCCGAGCAGCCATCCGCGATGGGAG.....TCCAAA 896
655 uIleThrGlnThrProSerAlaAlaProGluValTyrThrArgSerSers 672
897 CCGCAATGGCGCAGCAGCATAGAACCCCTCAGCAATATCTTACGGCAG 946
672 eThrMetProGluThrAla.....Gln 679
947 TCATCCCGCTCAAGAGGATTTGGAGCTTTCGGGAAATACGGCTTGGGC 996
680 SerThrPro.....Le 683
997 GGCATACGGCAGCATCTCTCAAGCGGTGCGAAGGCGGAGATGCGATT 1046
683 uAlaSerGlnAsnProThrSerSerGlyThrGlnThrHisAsnThrGluP 700
1047 GCGGAAGGGAATCCGCGCTCAGCGCAACATTTTCCGCGATGGGCGATAG 1096
700 rGlnTyr.....ThyTyrProValGlnThrThr 708
1097 CCAAAATACC.....CGTCCCTTACCATTCGCGA 1125
709 rProHisThrGlnLysLeuTyrThrGlnuAsnLysThrLeuSerPheProth 725
1126 AATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAGAAACATCACCTC 1175
725 rValValSerGluPheHisGluMetSerThrAlaGlnSerGlnThrPro. 741
1176 CTCACCGGCGCGCGCTCAACGGGAAGAAATGGAACCTGGCAACAAAC 1225
742 Leu.....LeuAspValLysIleValGlu..ValLysPheSerAsn... 754
1226 GCGACCGGAGAGCAAGAGTCCGTTTACGCGTAAGGTTTCCGAATTT 1275
755AspGlyGluValThrAlaThrCys 762
1276 GAAAAAGAGTAAAA.....TAGATACGAGAAATTAATACCGTGTACC 1319
763 ValSerThrValLysSerProTyrArgValGluThrAsnTrpLysValas 779
1320 ACAAGTAATCTATAGATGAACCGCTTTTAAAT...CCTAAAGGTTCTG 1366
779 rLeuValAspValMetAspGluLysSerGlyLysSerProAlaGlyValP 796
1367 TCGGATCGGCTCATCTTG.....TCTATACTGCGAGAAATTCATATAC 1410
796 hAsnSerAsnGluLysTrpGlnLysGlnLeuTyrTyrArgValThrAsp 812
1411 GCAAAATTACCAAGGCA..... 1428

813 GlyArgThrSerValGlnLeuMetCysLeuSerCysThrSerHisSerP 829
1429GTAAGATGAGATATA 1444
829 OGluProTyrCysLeuPheAspThrSerLeuIleAlaArgLysAspI 846
1445 TCCACCTAAATAATTAATCTTCCTTACAGCACC..... 1476
846 LeAlaProGluLeuTyrPheThrSerAspProGlnThrAlaTyrCysThr 862
1477CTACCAAAAGGAGCTAATATAGATATTTGGATTAATTTGGTAA 1520
863 IleThrLeuPro.....SerGlyValValProArgPhe..... 873
1521 TGAATG.....ACTA 1531
874 .GluTrpSerLeuAsnAsnValSerLeuProGluTyrLeuThrAlaThr 890
1532 AAGTCCATCAAGAACTAAAGTCAAGAAATTTGAATGGATGTTCAATTTG 1581
890 hrValValSerHisThrAlaGlyGlnSerThrValTTP..... 902
1582 TCTAAACAGGAGAGAGCAACTTGGATGGGCTAGTAGGATGTAAGCA 1631
903 LysSerSerAlaArgAlaGlyGluAlaTrpIleSerGlyArgGlyGlyAs 919
1632 TTTA 1635
919 nIle 920
seq_name: /c9n2_6/prodata/1/laa/6B_COMB.pep:US-09-103-429A-3
seq documentation block:
; Sequence 3, Application US/09103429A
; Patent No. 6187558
; GENERAL INFORMATION:
; APPLICANT: Granados, Robert R
; TITLE OF INVENTION: A No. 6187558el Invertebrate Intestinal Mucin
; TITLE OF INVENTION: CDNA and Related Products and Methods
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Pinnisi & Michaels, P.C.
; STREET: 118 No. 6187558th Tloga
; CITY: Ithaca
; STATE: NY
; COUNTRY: USA
; ZIP: 14850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/103,429A
; FILING DATE: 24-JUN-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Michaels, Christopher A
; REGISTRATION NUMBER: 34,390
; REFERENCE/DOCKET NUMBER: BTI-39
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (607) 256-2000
; TELEFAX: (607) 256-3628
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 786 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO

Align seg 1/1 to: US-08-728-323A-2 from: 1 to: 1162

```

284 CCGATACGGGC.....ACGAATCCATCCCCCTCGAC 318
    ||||| |||
63 ProleProGlySerProThrValPheThrSerGlyLeuProAlaPheVal 79
    :||| :|||
319 AACCATGCCCTACATCCGATTCTG.....ATGAAGCCGGTAG 356
    :||| :|||
79 LserSerProThrLeuProValAlaProLeuProSerProAlaProAlaT 96
    :||| :|||
357 TCCCGTTGACGATTACGCTTACCGCATTCATTGGACGATACGAAC 406
    :||| :|||
96 hrProLeuProProProAlaLeuLeuProProValThrThrSerSer 112
    :||| :|||
407 ACCATCCCGCCGACGCTATGACGGCCACAGGGCCGGCTATCCCGCT 456
    ||||| |||
113 ProleProProSer.....His..... 118
    :||| :|||
457 CCGAAAGCCGCGAGGATATATACAGCTACGACATAAAGCGTTGCCCA 506
    ||| :|||
119 .ProValSerProGly.....ThrThrAspThrHisSerProSerProA 133
    :||| :|||
507 AAATATCCCGCTCAACCTGACGACACCGACACCGGACGACGAGCTG 556
    :||| :|||
133 Laleu.....ProProThrGlnSerProGlnSerSerGln 144
    :||| :|||
557 TCGACGTTTCACATACCGGAGTAGTATGCTGACGACGAGGAGCGAC 606
    :||| :|||
145 ArgProProLeuSerSerProThrGly.....ArgProAspSerSerTh 159
    :||| :|||
607 GGATTCAAACGCGCCACCC.....GATACAGCCCGAGCTGACGACATC 650
    :||| :|||
159 rProMetArgProProProSerGlnGlnThrThrProProHisSerProT 176
    :||| :|||
651 GGGCATGCCGCG.....AAGCTTCAACGGCAGCTGCGAGTA 688
    :||| :|||
176 hrThrProProProGluProProSerLysSerSerProAspSerLeuAla 192
    :||| :|||
689 TCGTCAAAAACATCATCGCGCGGAGGAAATGTGCGCGGAGCGCAT 738
    ||| |||
193 ProSerThrLeuArgSerLeuArgLysArgArgLeuSerSerProGlnG 209
    :||| :|||
739 GCGGTGACAGGCTATTAAGCAGAGGCTCAACATTG..CTGTATGACCG 785
    :||| :|||
209 yPro.....SerThrLeuAsnProIleCysGln 219
    :||| :|||
786 CTGGGCTGCTTCCACGGAACAAAGATGGCGGCATCAACGATTGG 835
    :||| :|||
219 erProProValSerProProArgCysAspPheAlaAsnArgSerValTyr 235
    :||| :|||
836 CAGATATGGCGCACTCAAAAGACTATGCGGACGACCATCGCGGATTGG 885
    ||| :|||
236 ProProThrAlaThr.....GluSerProIleTyrValG 247
    :||| :|||
886 GCAGTCCAAAAC...CCAATGCCGACAAAGCATAGAACCGCTGCAGAA 932
    :||| :|||
247 ySerSerSerAspGlyAspThrProProArgGlnProProThrSerProI 264
    :||| :|||
933 TATCTTACGCGCAGTCATCCCGCTCAAAAGGATGGAGCTGTCCGGGAA 982
    :||| :|||
264 leSerIleGlySerSerSerProSerGlnGly.....SerTyrGlyAsp 278
    :||| :|||
983 AATACGGCT 991
    :||| :|||
279 AspThrAla 281
    :||| :|||

```

seq_name: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:US-08-628-829-4

seq_documentation_block:

; Sequence 4, Application US/08628829A
; Patent No. 6333170
; GENERAL INFORMATION:

```

; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To Exter
; FILE REFERENCE: CPI-004VPC3
; CURRENT APPLICATION NUMBER: US/08/628,829A
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: 08/440,421
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/323,460
; EARLIER FILING DATE: 1994-10-14
; EARLIER APPLICATION NUMBER: 08/049,254
; EARLIER FILING DATE: 1993-05-15
; EARLIER APPLICATION NUMBER: 08/410,602
; EARLIER FILING DATE: 1995-04-24
; EARLIER APPLICATION NUMBER: 08/472,934
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1593
; TYPE: PRT
; ORGANISM: Mus musculus
; US-08-628-829-4

```

alignment_scores: Quality: 119.50 Length: 495
Ratio: 0.595 Gaps: 29
Percent Similarity: 40.606 Percent Identity: 23.636

alignment_block:

US-09-303-518D-465 x US-08-628-829-4 ..

Align seg 1/1 to: US-08-628-829-4 from: 1 to: 1593

```

138 GAATACCACTATTCGCGACGAGGGGAACT..... 170
    ||||| |||
1 GluLeuProProSerArgCysProAlaArgSerArgThrThrPargArg 17
    :||| :|||
171 .....TGCGAGCGCGAGGCTCATATCGGATTGGGAAACATCAAAAGC 213
    ||| |||||
17 pGluGlyCysSerAlaGlnArg.....P 25
    :||| :|||
214 CATGATTGGGCAACCTGTTTCATCCAGCAGCGGCCATTAAAGAAATAT 263
    :||| :|||
25 roAlaProGlySerPro..... 30
    :||| :|||
264 CGGCTACATGTCCGCTTTCCGATCAGGCGACAGAGTCATTCCCTT 313
    :||| :|||
31 .....HisCysProCys..... 34
    :||| :|||
314 TCGACACCAATGCCCTCACATTCCGATTCTGTATGAAGCCGGTAGCCGTT 363
    ||| :|||
35 ....ProProThrCysAlaAlaArg..... 41
    :||| :|||
364 GACGATTACAGCTTACCGCATCCATTGGGAGCATACGAACACATCC 413
    :||| :|||
42 .....ProProAlaAlaProArgSerArgAlaProAla 52
    :||| :|||
414 ...CGCCGAGCGTATGACGCGCCACAGGGCGGCGCTATCCGCTCCA 460
    :||| :|||
53 GlyArgArgGlyProAlaAlaArgAlaArgAlaLeuGlySerSerAl 69
    :||| :|||
461 AAGCGCGAGGATATATACAGCTACGACATMAAAGGCTGCCCAAT 510
    ||| :|||
69 argProProThrArgProProLeuArgProProProAlaLeuSerProp 86
    :||| :|||
511 ATCCGCTCAACCTGACGACGACGACGACGACGACGACGCTTGTGCA 560
    :||| :|||
86 roProProSerPro.....AlaGlyThrSerGluCysSer 97
    :||| :|||
561 CCGTTTCCAAATACCGGTAGTAGCTGACGACGAAGAGTAGGACAC... 606
    :||| :|||
98 Pro.....Arg.GluLysMetAlaAlaAlaGlyAspArg 110
    :||| :|||

```

```

607 .....GATTCAAAGCGCCAGCATACAGCCCGAGCTGAC 645
110 LaserSerSerGlyPheProGlyAlaAlaAlaAlaSerProGlyAlaGly 126
646 AGATCGGGCAATGCCCGCAGCTTTCAGCGCATGAGATATACGCA 695
127 GlyGlyGlyGlyGlyGlyAlaLeuGlnGlySerGlyAlaProAla 143
696 AAACATCATCGCGCGCAGGAAATTCGGCGCGAGCGGATGCCGTC 745
143 aglyAlaAlaGlyLeuLeuArgGlyProGlySerAlaGlyProSerAla 160
746 AGGTATATAGGAGGCTCAACATCTGTTATGACGCGCTGG... 790
160 rg...ThnglyGlyGlyGlyThrcysAlaLysCysGlyValItrpSerTr 175
791 .....GTCTGCTTTCACCGCAAAACAGATGCG 818
175 PThrSerCysArgSerSerArgSerSerProProArgArgPro 191
819 GCGCATCAGCATTTGGCAGATATGGCGCACTCAAGACTATGCCGAG 868
192 AlaHis...LeuLeuProValAlaGlyAlaArgGlyArgGlyCysArgSe 207
869 CAGCCAT.....CCGCGATTGGCAGTCCAAAACCCCAATGCCGCA 912
207 rGluSerLeuProAlaArgAlaGlyProProProProGlyAlaAlaSer 224
913 GCGATAGAGCGCTCAGCAATCTTTCAGCGCATCCCGCTCAAGG 962
224 rg.....CysGlySerHis..... 228
963 GATTGAGCTGTTCGGGAAATACGCTTGGCGGCGCATACGCGCATC 1012
229 .....SerAlaGlyLeuAlaAlaAlaArgSerGlyAlaAr 241
1013 CTGTCAAGCGGTGCGAGATGGG...GAGATCGCATGCGCAAGGAA 1059
241 gSerProAlaGlyAlaGluProProSerAlaAlaAlaProSerGlyArg 257
1060 TCCGCGCTCAGCGCAATTTTCCGATGGCGCATCCCAATACCCGTC 1109
258 GluMetGluSerGlyThrLeuGlyGlyLeuHisLysMetGluArg 274
1110 CCTTACCATTCGCCGAATATCGGTCAAACTTGAGCAGCGTTAGCGCA 1159
274 gPro.....GluGluArgMetIle 281
1160 AAGAAACATCACCTCTCAACCGTCCCGCAACGGAAGATGTG 1209
281 rGlyLysLeuLysAlaThrCysMetProAlaTrpLysGluTrpLeu 297
1210 AAAGTGGCAACAAAGCCGACCGAAGACCAAGTCCGTTT... 1251
298 GluArgArgAsnArgArgGlyProValValLysProIleLeu 314
1252 .....GACGGTAAAGGTTTCCGAATTTGAAAAAGAGCTAAATACGATA 1297
314 sglyAspGlySerGlyValAsnAsnLeuAlaGlyProGlnGlyGlu 331
1298 CGAGAAATTAATACGCTGTACCAACAGTGAATCTATAGTGAACCGCTC 1347
331 lyGlnAlaGlySerAlaAlaProLa..... 339
1348 TTTATCTCTAAAGT.....TCGTGCGATCGCTATCTTG 1385
340 .....ProlysglyArgArgSerProSerProGlySerSerProSerG 354
1386 G...TCTATA.....ACTGCAGCAATTCAPAGCGCA 1414
354 yArgSerValLysProGluSerProGlyValArgGlyValSerP 371

```

```

1415 AATTACCAAGCGAA...GTAGATCATGATATATCCACTAAATATAC 1461
371 roValProGlnSerGlyArgIle.....ThrProProArgArg... 384
1462 TCTCTTCAGACCGCTACCAAAAGGA 1488
385 .....AlaProSerProAspGly 390

```

seq_name: /cgn2_6/plodata/1/iaa/6A_COMB.pep:US-08-899-437-2

seq_documentation_block:

; Sequence 2, Application US/08899437

; Patent No. 6121415

; GENERAL INFORMATION:

; APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESS: Genentech, Inc.

; STREET: 1 DNA Way

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/899,437

; FILING DATE: 24-Jul-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Conley, Delidre L.

; REGISTRATION NUMBER: 36,487

; REFERENCE/DOCKET NUMBER: P1084R1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650/225-2066

; TELEFAX: 650/952-9881

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 713 amino acids

; TYPE: Amino Acid

; TOPOLOGY: Linear

; FEATURE:

; NAME/KEY: Mouse NRG3 (mnRG3)/amino acid seq.

; LOCATION: 1-713

; IDENTIFICATION METHOD:

; OTHER INFORMATION:

; US-08-899-437-2

alignment_scores:

Quality: 118.50 Length: 554

Ratio: 0.517 Gaps: 30

Percent Similarity: 41.336 Percent Identity: 23.105

alignment_block:

US-09-303-518d-465 x US-08-899-437-2 ..

Align seg 1/1 to: US-08-899-437-2 from: 1 to: 713

```

275 TCCGCTTTCCG...ATCAGCGGCAGCAAGTCATTCCTCCCTTGACAC 321
119 SerSerPheProLysAlaMetGluThrThrThrThrThrSerThrTh 135
322 CATGCTCATCATTCGATTCTGATGAGCGGATGTCGGTTGACGAT 371
135 rSerProAlaThrProSerAlaGlyGlyAlaAlaSerSerArgThrPro 152
372 CAGCTTACCGCATTCATTTGGAGCGATACGAAACACATCCCGCGAG 421

```

```

416  erg1usSerLeuMetIysSerHisValHisLeuGlnHisnTrpIysSerIysAla 432
1084  GATGGCGGCAT.....ACGCCAATA 1103
      ||| |||||
433  Asp.ArgHisProValThrAlaLeuGlnIysIleMetGlnSerPheS 449
1104  CCGCTCCCTTACCATTCCCGAATATCCGGTTCACAACTGGAGCAGCGTT 1153
      :: ||| :::::::::::::::::::: ::|||
449  eraIarProGlnSerPheProGlnValThrSerProAsnArgIleGln 465
1154  ACGGCAAAAGAAACATCCACTCTCTCA.....CCGTGCCCGCGTCA 1194
      ||| ||||| ||| ::|||::
466  ProIleIysHisSerProGlnIleArgSerGlyMetLeuHisArgS 482
1195  ACGGAAAGATGTGAAACACGTGCACAAACAAAGCCACCAGAACCAAT 1244
      ::||| ||| ::||| |||::
482  nTrpPheArg.....ArgAlaProProSerProValS 493
1245  GCGCTTTG...ACGGTAAAGGTTTCCGATTTTGAAAAAGCATTAAT 1291
      ||||| ::| |||
493  eraIArgLeuGlyIleValIcIY..... 500
1292  ACGATACAGAAATTAATACCGCTGACACACACTGAATCTATAGATGA 1341
501  .....ProAlaTyGlnGln..... 505
1342  CCGCTCTTATATCTAAAGTCTGTGCGATCGCGCTCATTTGGTGTAT 1391
      |||::|||
506  .....LeuGlnGlnSerArgIle..... 511
1392  AACTGCCAGATTCATATACGCCAAATATTCACAGGC...AAGGTAGATCA 1438
512  ....ProAsnGlnAspThrIleProCysIleGlnIleGlnValArgIys 526
1439  GATATATCCCACTTAATAATTTACTCTCTCCACACCGCTACCAAAAGCA 1488
      ||||| ||||| ||| |||
527  ThrIleSerHisLeuProIleGlnIleTrpCysValGlnArgProLeuS 543
1489  CCTAATTAATGATATTGGATTAATTTGGTAATGTAATGCATGAAGCTCC 1538
543  PLeu.....LysTyGValSerAsnGlyLeuArgThrG 554
1539  ATCAA.....GACTAAAGTCAAGATTTGATGGG 1570
      ::||| ::||| ::|||
554  GlnGlnAsnAlaSerIleAsnMetGlnLeuProSerArgGlnThrAsnPro 570
1571  ATGTTCATTTGCTTAATAACAGAGACAGACAACTGGATGCGCTATAGG 1620
      ||||| ||| |||::: |||||
571  TyrPheAsnSerLeuAspGlnIysAsp.....LeuValG 582
1621  GATGATGAC 1630
      | ::|||
582  yTyrLeuSer 585
seq_name: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:US-09-126-121-2
seq_documentation_block:
  : Sequence 2, Application US/09126121
  : Patent No. 6252051
  : GENERAL INFORMATION:
  : APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
  : TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
  : NUMBER OF SEQUENCES: 23
  : CORRESPONDENCE ADDRESS:
  : ADDRESSEE: Genentech, Inc.
  : STREET: 1 DNA Way
  : CITY: South San Francisco
  : STATE: California
  : COUNTRY: USA
  : ZIP: 94080
  : COMPUTER READABLE FORM:
  : MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126.121
FILING DATE: 30-Jul-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: PI084R1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: Mouse NR3 (mNR3)/amino acid seq.
LOCATION: 1-713
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-09-126-121-2
```

```
alignment_scores:
Quality: 118.50      Length: 554
Ratio: 0.517         Gaps: 30
Percent Similarity: 41.336      Percent Identity: 23.105
```

alignment_block:

US-09-303-518D-465 x US-09-126-121-2 ..

Align seg 1/1 to: US-09-126-121-2 from: 1 to: 713

```
275 TCCGCTTTTCG...ATCAGGGGACGAAATCCATTCCTCCCTTGACAAAC 321
119 SerSerPheProLysAlaMetGluThrThrThrThrThrThrSerThr 135
322 CAGCCGTCACATTCGATTCGTATGAAGCGGAGTCCGCTGACGAGAT 371
135 rSerProAlaThrProSerAlaGlyAlaAlaSerSerArgThrProA 152
372 CAGCCCTTACCGCATTCATGGGACGATAGCAACATCCCGCGAGC 421
152 snArgLieserThrArgLeuThrThrLleThrArgAla.....ProThr 166
422 GCATGACGGGCGACAGGGCGGCGGTATCCCGCTCCCAAGGCGCGAG 471
167 ArgPheProGlyHisArg.....ValProIle..... 175
472 GATATATACAGCTACGACATAAAGGCGTTGCCAAATATACGCGCTCA 521
176 .....ArgAlaSerProArg.....SerThr 183
522 CCGTACGACAAACCGACGACGACGACGCTGTGACGCTTTCACA 571
183 hrrAlaArgsnThrAlaAlaProProThrValLeuSerThr.....Thr 197
572 ATACCGGATGATGCTGACGACGAAGAGTAGGCGACGAGATCAACGCGCG 621
198 AlaProPhePhe.....SerSerSerThrProGlySerArgPr 210
622 ACCG..... 625
210 orPheMetProGlyAlaProSerThrGlnAlaMetProSerThrProThrA 227
626 .....GATACAGCCCGAGCTGACGA 646
227 AlaLysGlnAlaThrSerSerGlyLeuHisAspSerThrProSerThrPrh 243
```

```
646 ..... 646
244 LeuSerProPheGlnAspAlaAlaAlaSerSerSerProSer 260
647 .....GATCGGCAATGCGCGGACGCTTTCACAGCGACTGACAGAT 689
260 rThrSerSerThrThrThrThrProGluThrSerThrSerProLysPhe 277
690 CGTCAAAAACATCATCGCGCGGCGACGAGAAATGTGCGCGCAGCGCAT 739
277 IsThrThrThrThrSerThrGluArg.....SerGlnHisPheLys 290
740 CCGTCAGGCTATAGCAAGGCTCAACATTGCTGTATGTC...ACGGC 786
291 ProCysArg.....AspLysAspLeuAlaLysLysLeuAsnAs 303
787 TTGGGTGCTGCTTCCACCGAAACA.....AGATGC 818
303 pGlyGluCysPheValIleGluThrLeuThrGlySerHisLysHisCysA 320
819 GCGCATCAACGATTTGGCAGATATGCGCGCACTCAAGACTATGCGCGCAG 868
320 rGysLysGluGlyTyrglnGlyAlaArgCysAspGlnPheLeuProLys 336
869 CAGCCATCCGCGATTTGGGACAGTCCAAAACCCCAATGCGCGCAAA.GGCAT 917
337 ThrAspSerIleLeuSerAsp.....ProThrAspHisLeuGlyI 350
918 AGAA.....GCCGTCA 928
350 eGluPheMetGluSerGluAspValTyrglnArgGlnValLeuSerLies 367
929 GCAATATCTTTACGGACGATCAATCCCGCAAGAGATTTGAGCGGT... 975
367 eCysIleIlePheGlyIleValIleValIleGlyMetPheCysAlaIlePhe 383
976 .....CGGGAAATAATACGGCTTGCGGCGCATCACGGCACATCTGTCAA 1019
384 TyrPheLysSerLysLysGlnAlaLysGlnIleGlnGlnHis...LeuLys 399
1020 GCGGTCGACGATGGGCGGAGATCGCATTCGCGAAAGGAAATCCCGCTCA 1069
399 sGluSerGlnAsnGlyLysAsnTySerLeuLysAlaSerSerThrLys 416
1070 GCGACAATTT.....GCC 1083
416 eGluSerLeuMetLysSerHisValHisLeuGlnAsnTySerLysAla 432
1084 GATCGGCGCAT.....ACGCCAATA 1103
433 Asp.ArgHisProValThrAlaLeuGluLysIleMetGluSerSerPhe 449
1104 CCGGTCGCCCTTACCATTCGCGAAATATCCGTTCAACTTGGACAGCGTT 1153
449 eAlaProGlnSerPheProGluValThrSerProAspArgGlySerGln 465
1154 ACGGCAAAAGAAACATCAGCTCTCA.....CCGTGCGCGCGTCA 1194
466 ProLleLysHisHisSerProGlyGlnArgSerGlyMetLeuHisArgAs 482
1195 AACGGAAGAAGTGTGAACCTGGCAACAAAGCCACCGGACGAAAGT 1244
482 nThrPheArg.....ArgAlaProProSerProArg 493
1245 GCGCTTTG...ACGTAAGGCTTTCGGAATTTTGAAGAACGCTAAAT 1291
493 eArgLeuGlyGlyIleValGly..... 500
1292 ACGATACGAGAAATTAACCGCTGTACCAAGTAAGTCTATAGATGA 1341
501 .....ProAlaTyrglnGln..... 505
1342 CCGGTCCTTAATCCTAAAGGTTCTGTCGATCGGCTCATTTCTGTGTAT 1391
```

```
506 .....leuGluGluSerArgIle..... 511
1392 AACTCCAGATTCATCAATTCAGCAAAATTCACAGGC...AAGGTAGAATCA 1438
512 .....ProAspGlnAspThrIleProCysGlnGlyIleGluValArgLys 526
1439 GATATATCCACCTAAATAATTACTCTTCACACCGCTACCAAAAGA 1488
527 ThrIleSerHisLeuProIleGlnLeuThrProCysValGlnArgProLeuAs 543
1489 CCTAATATAGGATATTGGATAATTGTGAATGATGACTAAAGTCC 1538
543 Pleu.....LysTyrValSerAsnGlyLeuArgThrG 554
1539 ATCA.....GAATAAGGTCAAGAATTGATGGG 1570
554 InGlnAsnAlaSerIleAsnMetGlnLeuProSerArgGlnThrAspPro 570
1571 ATGTTCAATTGTCTAAACAGCAAGACGACTTGATGGCTAGTAG 1620
571 TyrPheAsnSerLeuAspGlnLysAsp.....LeuValGln 582
1621 GATCGTAGC 1630
582 YTYrLeuSer 585
```

seq_name: /cgn2.6/ptodata/1/1aa/6B_COMB.pep:US-09-413-814-79

```
seq_documentation_block:
; Sequence 79, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bioecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hoffe, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; EARLIER FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 1213
; TYPE: PRF
; ORGANISM: Sorangium cellulosum
US-09-413-814-79
```

alignment_scores:
Quality: 116.50 Length: 324
Ratio: 0.910 Gaps: 15
Percent Similarity: 39.506 Percent Identity: 26.235

alignment_block:

US-09-303-518D-465 x US-09-413-814-79 ..

Align seg 1/1 to: US-09-413-814-79 from: 1 to: 1213

```
105 GGTCTCGACCGTCAGCATTTCCGAACCCGAGCGAAATACCACTATTTCG 154
||||| ||||| ||||| ||||| |||||
351 GtYAlaValProGlnAlaProHisAspArgArgAlaAlaValGlnAspArg 367
```

```
155 GCAGAGGGGGGAACTGCCGACGCGACGCGTACATTCGATTGGGAAC 204
||||| ||||| ||||| ||||| |||||
367 gArgrProGlnProLeuLeuArg...ArgArgAspHisArgAlaAlaArg 383
205 ATACAAGCCATTCAGTTGGCAACCTGTTCATCCAGCAGCGGGCCATTAA 254
||||| ||||| ||||| ||||| |||||
383 lAnHArgProSerGlnGlnAspProArgLeuPro..... 394
255 AGGAATATACGGCTACATTGTCCGCTTTCCGATACGGGACGACGATCC 304
||||| ||||| ||||| ||||| |||||
395 .....HisArgAlaArg..... 398
305 ATTCGCCCTTCGACAACATGCGCTCACATTCGATTGTGAAGAACCGGT 354
398 ..... 398
355 AGTCCCGTTGACGATTCAGCTTACCGCATTCATG..... 392
||||| ||||| ||||| ||||| |||||
399 .....ArgAspArgGlnArgProArgAlaAlaProLeuGlnArgAlaGly 413
393 .....GACGCGAT 400
414 GtYArgGlnGlyLysArgSerValGlnArgGlnAlaProGlnYArgVa 430
||||| ||||| ||||| ||||| |||||
401 ACGACACCATCCCGCGCGCTATGACGGGCGACAGGGCGCGCTAT 450
430 lArgArgArgArgArgArgArgArgArgArgArgArgArgArgArg 444
451 CCGCGTCCCAAGCGCGCGAGATATATACAGCTACGACATTAAGGCGT 500
||||| ||||| ||||| ||||| |||||
445 ...ArgProGlnAlaAlaGlnGlyValHisAspProArgGlnYArgArg 460
501 TGCCCAAAATATCCGCTTACCTGACGCGACACCGCGACCGGACAC 550
||||| ||||| ||||| ||||| |||||
461 ProProGlnYAlaAlaAlaGlnArgGlnArgGlnGlnYArgProArg..... 475
551 GCGTTCGACCGCTTCCACATACCGGTATGCTGACGCGACGAGTA 600
||||| ||||| ||||| ||||| |||||
476 .....ArgProAlaGlnYSerArgArgGlnYArgAlaAlaArgArgGly 490
601 GCGCA.....CGGATTCAAACGCGCGACCGGATACGCGCGAGCT 641
||||| ||||| ||||| ||||| |||||
490 rGArgProAlaAspArgArgArgArgAlaAlaHisArgLeuGlnYAlaArgGly 506
642 GACAGATGGGCAATGCCGCGGAAAGCTTTCACAGGCACTGCAAGATTCG 691
||||| ||||| ||||| ||||| |||||
507 GtYLeuAlaGlnYArgGlnYArgArg.....Arg 515
692 TCAAAACATCATCGCGCGCGAGAGAAATGTGCGCGACGCGATGCC 741
||||| ||||| ||||| ||||| |||||
515 gAspArgGlnProLeuArgArgAlaArgLeuHisLeuAlaAlaProArgAla 532
742 GTGACAGGTATTAAGCGAAGCTCAAAATGCTGTATGACAGCGCTTGGG 791
||||| ||||| ||||| ||||| |||||
532 rGAlaAlaProAlaAlaArgAla..... 539
792 TCTGCTTTCACCGAAGAACAGATGGCGCGCATCAAGATTGGCAGATA 841
||||| ||||| ||||| ||||| |||||
540 .....HisArg.....GtYAlaArgAlaArgArgGlnYArgArg 550
842 TGCGCAACTCAAAAGCATATGCCGCGACGACGATCCGATGGCGACATC 891
||||| ||||| ||||| ||||| |||||
550 gGtYAlaGlnYArgValAlaAspArgPro.....ValP 561
892 CAAAACCCCAATGCGCGACAGCATAGAACGCTGACGACATATCTTTAC 941
||||| ||||| ||||| ||||| |||||
561 rValPProAsp.....HisArgAlaThr..... 568
942 GCGAGTCATCCCGGTCAAGGATGGATGGAGCTGTGGGGAAATACGGCT 991
||||| ||||| ||||| ||||| |||||
569 GtYAlaAlaAlaArgArgGlnYAsp.....GtYGlnGlnYArgAl 581
```

992 TGGCGGCATCAGCGCACATCC 1013
|||||:|||||:
581 aglyArgrArgrGlyAlaala 588

seq_name: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:US-08-819-013-1

seq_documentation_block:

Sequence 1, Application US/08819013
Patent No. 5994522
GENERAL INFORMATION:
APPLICANT: Chan, Andrew C.
TITLE OF INVENTION: BLNK PROTEINS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
FILING DATE: 17-MAR-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/798,322
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-64383-1/RT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-819-013-1

alignment_scores:

Quality: 116.00 Length: 237
Ratio: 1.036 Gaps: 9
Percent Similarity: 47.257 Percent Identity: 24.895

alignment_block:

US-09-303-518D-465 x US-08-819-013-1 ..

Align seg 1/1 to: US-08-819-013-1 from: 1 to: 456

236 TCACGACGAGCGCCATTAAAGAAATCGGCATATGTCGCTTTCC 285
|||||:|||||:
124 SerSerInArHisSerProPheSerIyThIeuProSerIySpr 140
286 GATCAGCGGACGACGATCCCTTCGACACACATCCCTCACATTC 335
|||||:|||||:
140 OSeITriProSerIeIuIyAlaArgIeuThIeuProAlaIeuT 157
336 CGATTGATGACGAGCGGTAGTCCG.....TTGACG 367
|||||:|||||:
157 hralaIeuGInIySProlInValIProProlYSProlYsgIyIeuIeuGlu 173
368 GATTACGCC...TTTACCGCATCCATT...GGGACGAGTACGAACACCAT 411

|||||:|||||:
174 AspGluAlaAspIyArValIProValGluAspAsnAspGluAsnIyTIL 190
412 CCGCGCGAGCGGCTATGACGCGGCACAGCGCGCGCTATCCGCTCCCAA 461
|||||:|||||:
190 eHisProIthr.....GluSerSerProProIgluL 202

462 AGCGCGAGGATATATACAGCTACGACATATAAGCGCTTGCCCAATA 511
|||||:|||||:
202 yAla.....PrometVal 206

512 TCGGCTCACTGACCGACACACCGACCGACGACAAAGCTGTGAC 561
|||||:|||||:
207 AsnArgSerThIySProlAsnSerSerThIyProIaSerProIglYTh 223

562 CGTTTCCAAATACCGGTGTGTGTCGACGAGAGAGAGGACGAGT 611
|||||:|||||:
223 rAlaSerGly.....ArgAsnSerGlyAlaIatIrgIuT 234

612 CAACGCGCCACCGATACAGCCCGAGCTGACAGATCGGCA..... 655
|||||:|||||:
234 hIySerProProIaIaIProSerProIeuProIaIaIyLys 250

656ATGCCCGGAGCTTCAACGCGACTGCAGATATGCTCAAAAC 699
|||||:|||||:
251 IySProlThIyProIeuIyThIyProIaIaIaIaSerGInIaSnAl 267

700 ATCATCGGCGGCGGACGAGGAAATGTGCGCGCGACGATGCCGTACAG 749
|||||:|||||:
267 aSerSerValIySGluGluLys..... 274

750 TATAAGCAGAGGCTCAACATTGCTTATATGACGCGCTGTGCTTT 799
274 274

800 CCA...CCGAAACAAGATGCGCGCATCATCATTTGGCAGATATGCG 846
|||||:|||||:
275 ProIleProIaIaIArgHisArgIySerSerHisArgGInIaIaVa 291

847 CAACCTCAAGACTATGCGCGACGACGACCATCCGCTTGGGAGTCCAAA 896
|||||:|||||:
291 IGlSerProValIhneProIaIaIaIySGInIleHisGInIySProl 308

897 CCCCAATGCG 907
|||||:|||||:
308 IeProIeuPro 311

seq_name: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:US-08-793-824-2

seq_documentation_block:

Sequence 2, Application US/08793824
Patent No. 5981838
GENERAL INFORMATION:
APPLICANT: Simpson, Christine Lynn
APPLICANT: Giffard, Philip Morrison
APPLICANT: Jacques, Nicholas Anthony
TITLE OF INVENTION: Genetic Manipulation of Plants to
INCREASE STORAGED CARBOHYDRATES
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Griffith Hack & Co
STREET: Level 8, 168 Walker Street
CITY: No. 5981838th Sydney
STATE: New South Wales
COUNTRY: Australia
ZIP: 2060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,824


```

1165 AACATCACTTCCTCAACCGCGCCGCTCAACGGAAGAATGGAACCT 1214
      ::::::::::::::::::::
1120 Lysilethrglntrpseralalysrlypheasnlgysersnileglncl 1136
1215 GCGAAACAACGCCAC..... 1230
1136 ythrglyalargrtyvalleuglnaspanalathrasnlglnlyrphes 1153
1231 .....CCGAAGACCAAGTCCGCTTGAC 1254
1153 ervallysalaglyglntrpheleuprollysglnmettrngluilthr 1169
1255 GGTAAAGGTTTCCGAATTTGAAAAAGAGTAATACGATAGC.... 1299
1170 Glyserglypheargrvalglyaspsalvalglntrleuserilegl 1186
1300 .....AGATTAATACCGCTGACCAAGAGTAATCCATATGATG 1339
1186 yglytyrleualalalsasnthrphele...glnvalglyalaaasnlgln 1202
1340 AACCCGCTTTAAATCCTAAGGTTCTGCGATCGGCTCATCT.... 1383
1202 rprtyrtyrphesaplysasnlglyasmetvalthrclglnlglnval 1218
1384 .....TGGTCTATTAATGCCCAATTCATACGCAAAATT 1418
1219 Aspqllystysrtyrphelenuaspasnlglyleuglnleuargrhlsya 1235
1419 ACCAAGGCAA.....GTAGAAATCAGATATATCCCACTCAAAAT 1459
1235 lleuargrglnlglyseraspqllyhlsvallytyrtyraspprollysgly 1252
1460 ACTCTCTTCGACGCGCTACCAAAAGACCTAATGATGATTTGGAT 1509
1252 alglinalphe.....Asnlglyphetryrpsp 1260
1510 AAATTTGGT.....AATGAAG 1526
1261 Phealaglyproargrlnaspsalargtyrphesapqllyasnlglylme 1277
1527 GACTAAAGCCATCAAGAACTAAAGTCAAGATTTGATGGATGTTTC 1576
1277 ttryaraglyleuuhlsaspmetyrtyrtyrthrphearyrphesap... 1292
1577 AATTGTCTAAACAGCA 1593
1293 .....Gluysrthrgly 1296

seq_name: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:US-09-103-429A-4

seq_documentation_block:
; Sequence 4, Application US/09103429A
; Patent No. 6187558
; GENERAL INFORMATION:
; APPLICANT: Granados, Robert R
; TITLE OF INVENTION: A No. 6187558e1 Invertebrate Intestinal Mucin
; TITLE OF INVENTION: CDNA and Related Products and Methods
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Pinnisi & Michaels, P.C.
; STREET: 118 No. 6187558th Tl09a
; CITY: Ithaca
; STATE: NY
; COUNTRY: USA
; ZIP: 14850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/09/103,429A
FILING DATE: 24-JUN-1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Michaels, Christopher A
REGISTRATION NUMBER: 34,390
REFERENCE/DOCKET NUMBER: BTI-39
TELECOMMUNICATION INFORMATION:
TELEPHONE: (607) 256-2000
TELEFAX: (607) 256-3628
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 805 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Trichoplusia ni
TISSUE TYPE: peritrophic membrane
US-09-103-429A-4

```

```

alignment_scores:
Quality: 115.50      Length: 319
Ratio: 0.831        Gaps: 16
Percent Similarity: 43.574      Percent Identity: 23.824

alignment_block:
US-09-303-518D-465 x US-09-103-429A-4 ..

```

Align seg 1/1 to: US-09-103-429A-4 from: 1 to: 805

```

119 AGCATTTGAAACCG...ACGGGAATATACACCTATTTCGCGAGGCGG 165
      ::::::::::::::::::::
446 Thrilealargprocyproglyasnleuhlspheserproalathrgl 462
166 GAACCTGGCCGAGCGGCGATCATGCGATTTGGGAACATACAAACCA 215
      ::::::::::::::::::::
462 nserysgluser...ProvalthrAlaaglycysglvalaPheglucysa 478
216 TCAGTTGGCAACCTGTTTCATTCACAGCGCGCCATTAAAGAAATATCG 265
      ::::::::::::::::::::
478 spserAspsasnlglnrcysrthrserthrAlaAlaprothrAlaAlaprothr 494
266 CCTACATTGTCGCTTTTCGATCAGCGGCAAGATTCATCCCTTC 315
      ::::::::::::::::::::
495 AlaAlaprothrAlaAlapro...ThrAlaAlaprothrAlaAlaprose 510
316 GACAACCATGCTCCACAT.....CCGATTCGTGATGAAGCGGTAGTC 359
      ::::::::::::::::::::
510 rthvalvalproproalathrproproalathrAlaAlaprothrAlaAla 527
360 CGTTGACG.....GATTCAGCGCTTACCGCATTC 388
      ::::::::::::::::::::
527 rothrthrAlaAlaprothrProalaprothrAlaAlaprothrAlaAla 543
389 ATTGGGACG.....GATTCAGCAACCATTCGCGGAC 420
      ::::::::::::::::::::
544 prothrthrAlaAlaprogluserprothrthrvalthrvalproproth 560
421 GGCTATG...ACGGCCACAGCGGCGGTATCCG..... 454
      ::::::::::::::::::::
560 rAlaAlaprothrAlaAlaprothrthrAlaAlaAlaAlaAlaAlaAla 577
455 .....CTCCCAAGGCGCGA 469
      ::::::::::::::::::::
577 hrvalthrserAlaprothrAlaAlaAlaprothrAlaAlaAlaAlaAla 593
470 GGGATATATACAGCTACGACATAAAGCGCTGGCCCAAAATATCCGCTC 519
      ::::::::::::::::::::

```



```

594 ProthralaIaIaProthrThr.....AlaValProGluIleProthrth 608
520 A...ACCTGACGACAAACCGACACCGGACAAAGCGCTG.....T 557
608 ValThrSerProProthralaIaIaProthrThrIaIaIaProIaIaProIa 625
558 CGACCGCTTCACAAATACCGGTATGCTGACGACGACGAGTAGGAGG... 604
625 snThrThrValThrValProProthralaIaIaProthrThrIaIaIaPro 641
605 .....ACGGATTCAAACGCGCCACCGCATACGCCCGGAGCTGG... 643
642 AlaProsnThrThrValThrIaIaProProthralaIaIaProthrThrIa 658
644 .....ACAGATC 650
658 aAlaProIaIaProsnThrThrValThrValProProthralaIaIaProT 675
651 GGGCATGCGCGCGAAGCTT.....TCACGCGCATGCGAC 685
675 hrIaIaIaProProthrValaIaIaIaIaIaProsnThrThrIaIaIaPro 691
686 ATATGCTCAAAACATCATCGCGCGG..... 712
692 ValThrThrThrSerAlaProIaIaIaIaIaIaProGluAspAspIleAs 708
713 .....CAGGAAATTTGCGCGGACGCGCATGCCGTGCGAGG 749
708 pProIeProIaProsnAspProIleAsnProCysValGluGluCysAsnV 725
750 TATAGCGAAGGCT..... 763
725 aIleuProIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 741
764 .....CAACATTCGTGTATGACGCGCTTGCTGCTTCCACCGAA 807
742 AsnAsnGlnValIleuValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 758
808 AACACGA 814
758 rThLys 760

seq_name: /cgn2_6/prodata/1/laa/Backfiles1.pep:5268270-2
seq_documentation block:
; Patent No. 5268270
; APPLICANT: Meyer, Thomas F.; Halter, Roman; Pohlner, Johannes
; TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEINS USING GRAM
; NEGATIVE HOST CELLS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/171,872
; FILING DATE: 01-JUL-1987
; SEQ ID NO: 2
; LENGTH: 1507
5268270-2

alignment_scores:
  Quality: 114.50      Length: 580
  Ratio: 0.458         Gaps: 31
  Percent Similarity: 43.103  Percent Identity: 20.000

alignment_block:
US-09-303-518d-465 x 5268270-2 ..
Align seg 1/1 to: 5268270-2 from: 1 to: 1507

91 TCTTTATCCGCGAGCTTCGACCGTCGACATTCGAAACCGGAGAA 140
|||||
187 SerPheValArgIleGlyAlaGlyArgGlnLeuValTyrGluLysGlyVa 203
141 ATACCACTATTCCGCGAC...AGGGGGAACCTTCCGACGCGACGCGTGC 187

```

```

|||||
203 lTyrHisGlnGluGlyAsnGluLysGlyTyrAspLeuArgThrGluGlyL 220
188 ATATGGATGGGAAC.....ATACAAAGCATCACTGGCC 225
|||||
220 euIleGlyPheGlyAsnHisAsnLysGlnTyrSerAlaGluGluLys 236
226 AACCTGTTCATCCAGCAGCGCCCATTAAGGAATATCGCTCATGTGT 275
237 GlnAlaLeuSerGlnAspAlaLeuThr.....AsnTyrGlyValLeu. 250
276 CGGCTTTTCGATCAGCGGACGACGAATCCATTCCTTCGACAAACATG 325
250 ..... 250
326 CCTCATTTCCGATTCGTATGAAGCCGCTAGTCCGTTGACGATTCAGC 375
251 .....GlyAspSerGlySerProLeuPheAlaPheAsp 261
376 CTTTACCGCATTCATTTGGAGCGATACGAACACATCCCGCGACGCTA 425
262 LysGlnLysAsnGlnThrValPhe.....LeuGlyThrTyr 273
426 TCACGGGCCACAGGGCGGCTATCCGCTCCAAAGCGCGAGGATA 475
273 Asp.....TyrThrAlaGlyTyrGlyLysSerT 284
476 TATACGCTACGACATA.....AAAGGCTTGCCCAAAATTCGCGCTC 519
284 rPoleIleThrAsnIleTyrLysGluPheAlaIaIaIaIaIaIaIaIaIa 300
520 AACCTGACCGACACCGGACGACG.....GGACAAAGCGCTGT 557
301 His.....AspAsnAlaGlyThrValLysGluAsnGluGlnHis 313
558 CGACGCTTTCACAAATACCGGTATGCTGACGCGACGAGTAGGCGACG 607
314 .....HistPryLysThrGlyThr.....AsnSerHisIleGlySerT 327
608 GATTCAACGCGCGCGCATGACGCGCGGAGCTGGACGATCGGCGCAT 657
327 hrIaIaValArgLeuAlaAsnAsnGluGlyAspAlaAsnAsnGluGlnAsn 343
658 GCCCGCGAAGCTTTCAACGCGACATCGACATATGTCMAAACATCATCGG 707
344 ValThrPheGluAspAsnGlyThrIleuValIleAsnGlnIleAsnGln 360
708 CGCGCGACGAGAAATTTGCGCGCGACGCGAT...GCCGTGACAGGTATA 754
360 nGlyAlaGlyGlyLeuPheGluLysGlyAspTyrThrValLysGly.... 375
755 GCGAAGGCTCAACATTCGTGTATGACGCGCTTGCTGCTTCCACG 804
376 .....AlaAsnAsnAspIleThrTrpLeuGlyAlaIle..... 387
805 GAAACAGATGCGCGCATCAACGATTTGCGAGATTTGCGCAACTCAA 854
388 .....AspValAlaAspGlyLysValIaIaVala 396
855 AGACTATGCCGACGACGACATCCGCGATTTGGCAGTTCGCAAAACCCATG 904
396 l.....TrpGlnValLysAsnProIaIaIa 404
905 CCGCACAGCATAGAAAGCGTCAGCAATATCTTTACGAGAGCATCCCG 954
404 lY.....AspArgLeuAlaLysIleGlyLysGlyThrLeuGlu 416
955 GTCAAAGGATTTGAGCTTTCCGGGAAATACGCGTGGCGCGCATCAC 1004
417 lIleAsnGlyThrGlyValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 433
1005 G.....GCACATCTGTCAAGCGGTCGACG...ATG 1033

```


622 ACCC.....GATA 629
221 tProSerTrpProThrAlaAlaTyrAlaThrSerSerTyrLeuHisAspS 238
630 CAGCCCGAGCTGACAGATGGGCAATGCCCGAAGCTTTCAACGGCA 679
238 erThrProSerTrpThr.....LeuSerProPheGlnAspAlaAla 251
680 CTGACGATATCTCAAAACATCATCGCCCGG..... 712
252 SerSerSerSerSerSerSerSerAlaThrThrThrProGluTh 268
713CAGAGAAATGTCCG 728
268 rSerThrSerProLysPheHisThrThrThrThrThrGluArgSer 285
729 CGCAGGCGATCCGTCCAGGATTAAGCGAAGCTCAACATTCGTGTA 778
285 LuHisPheLysProCysArg.....AspLysAspLeuAlaTyr 297
779 TGC...ACGGCTTGGCTCTGCTTTCCACCGAACA..... 811
298 CysLeuAsnAspGlyGluCysPheValIleGluThrLeuThrGlySerH 314
812AGATGGCGCGCATCAACGATTTGGCAGATATGGCGCACTCAAGA 857
314 slySHisCysArgCysLysGluGlyTyrGlnGlyValArgCysAspGln 331
858 CTATGCCGACAGACCATCCGCGATTTGGGACATCCAAACCCCAATGCCG 907
331 heLeuProLysThrAspSerIleLeuSerAsp.....ProThrAsp 344
908 CAC..... 910
345 HisLeuGlyIleGluPheMetGluSerGluGluValTyrGlnArgGln 361
910 910
361 lleuSerIleSerCysIleIlePheGlyIleValIleValIleGlyMetPheC 378
911AAGCATAGAACCGCTCAGCAATA 934
378 ysaAlaIaPheTyrPheLysSerLysGlnAlaLysGlnIleGlnGlu 394
935 TCTTACGGCGCATGATCCCGCTCAAGGATGAGCTGTGGGGGAAA 984
395 GlnLeuLysValProGlnAsnGlyLysSerTyrSerLeuLysAlaSer 411
985 TACGGCTTGGCGCATCAGCGACATCTGTCAAGCGGTGCGAGATGG 1034
411 rThr.....MetAlaLysSerGluAsnLeu 420
1035 CGAGATGGCATTTGCCGAAGGAAAT.....CCGCG 1066
420 allySerHisValGlnLeuGlnAsnTyrSerLysValGluArgHisPro 436
1067 TCAGCGACAATTTGCCGATCGGCATACGCCAAATCCCGCTTAC 1116
437 ValThrAlaLeuGluLysMetMetGluSerPheValGlyProGln 453
1117 CATTCGCGAATATCCGTTCAACTGTGAGCAGCTTACGCAAGAAA 1166
453 rPheProGluValProSerProAspArgGlySerGlnSerValLysHisH 470
1167 CATCACTCTCTCAACCGTCCGCGCTCAACGAGAAAGATGTGAACGTG 1216
470 IsArgSerLeuSerSerCysSerProGlyGlnArgSerGlyMetLeu 486
1217 CAACCAACGCCACCGAAGA.....CCAAAGTCCCTTTG.. 1252
487 HisArgAsnAlaIaPheArgArgThrProProSerProArgSerArgLeuG 503
1253 .ACGCTAAGGTTTCCGATTTGAAAAAGACGTAAATAGATACGAG 1301

503 yGlyIleValGly..... 507
1302 AATTATACCGCTGTACCAAGATGATCTATGATGACCGGCTTTA 1351
508ProAlaTyrGlnGln..... 512
1352 ATCTAAAGTTCTGTGAGATGCGCTCATTCGTGCTATAATGCCAGA 1401
513LeuGluGluSerArgIle.....ProAs 520
1402 ATTCATACGCAAAATTACCAAGGC...AAGGTAGATCAGATATATCC 1448
520 pGlnAspThrIleProCysGlnGlyIleGluValArgLysThrIleSerH 537
1449 ACCTAAATTTACTCTC.....CTTCAGACCGCTAC 1480
537 IsLeuProIleGlnLeuTyrPysValGluArgProLeuAspLeuLysTyr 553
1481 CAAAGGACCTAATATGATGATTTGATTAATTTGTAATGATGACT 1530
554 SerSer.....SerGlyLe 558
1531 AAAG.....GTCCATCAGACTAAAGTCAAGAAAT 1562
558 ulsThrGlnArgAsnThrSerIleAsnMetGlnLeuProSerArgGluT 575
1563 TGAATGGGATGTTCAATTGTCTAAACAGAGAGAGACACTTGATGG 1612
575 hrAsnProTyrPheAsnSerLeuGlnGluLysAsp..... 586
1613 CTAGTAGGATGTATAGC 1630
587 LeuValGlyTyrSerSer 592
seq_name: /cgn2_6/ptodata/1/aa/6B_COMB.pep:US-09-126-121-6
seq_documentation_block:
; Sequence 6, Application US/09126121
; Patent No. 6252051
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
; TITLE OF INVENTION: ligands and Uses Therefor
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/126,121
; FILING DATE: 30-Jul-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Conley, Delidre L.
; REGISTRATION NUMBER: 36,487
; REFERENCE/DOCKET NUMBER: P1084R1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-2066
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 720 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear

```

FEATURE:
NAME/KEY:  INRG3B1 amino acid sequence
LOCATION:    1-720
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-09-126-121-6

```

```

alignment_scores:
Quality: 114.00      Length: 556
Ratio: 0.504         Gaps: 25
Percent Similarity: 40.647   Percent Identity: 20.863

```

alignment_block:

US-09-303-518D-465 x US-09-126-121-6

Align seg 1/1 to: US-09-126-121-6 from: 1 to: 720

```

275 TCCGCTTTCCG...ATCAGCGGCGAGAGTCCATTCCTCCGTCGACAAAC 321
    |||:|||||:|  |||:|  |||:|  |||:|  |||:|  |||:|
119 SerSerPheProLysAlaMetGluThrThrThrThrThrSerThrTh 135
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
322 CATGCCACATTCGATTCGTATGAAGCGGTAATCCGTTGAGGAT 371
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
135 rSerProAlaThrProSerAlaGlyAlaAlaSerSerArgThrPro 152
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
372 CAGCCTTTACCGCATCCATTTGGAGCGAGTAACAGAACATCCGCGCAG 421
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
152 snArgIleSerThrArgLeuThrThrIleThrArgAla....ProHn 166
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
422 GCATATGACGGCCACAGAGCGGCGGCTATCCGCTCCAAAGGCGGAG 471
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
167 ArgPheProLysIleSarValProIleArgAlaSerProArgSer... 181
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
472 GATATATACAGTACGACATAAAGCGTTGCCCAAAATATCCGCTCAA 521
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
182 .....ThrThrAlaArgAsnThrAlaAlaPro.....AlaArg 192
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
522 CTGACCGGACAGACCGGACGCGGACACGCTGTGACGCTTTCCACA 571
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
192 hrValProSerThrThrAlaProPhePheSerSerThrLeuGlySer 208
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
572 ATACCGGTGATGCTGACGAGCAAGAGTAGGCGAGCATTCAAAGCGCC 621
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
209 ArgProProVal.....ProGlyThrProSerThrGlnAlaMet 221
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
622 ACCG.....GATA 629
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
221 tProSerThrProThrAlaAlaAlaAlaThrSerSerThrLeuHisAsp 238
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
630 CAGCCCGGAGCTGGACAGATCGGGCAATCCCGCAAGCTTCAACGGCA 679
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
238 erThrProSerThrThr.....LeuSerProPheGlnAspAlaAla 251
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
680 CTCAGATATCGTCGCAAAACATCATCGGGCGG..... 712
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
252 SerSerSerSerSerSerSerSerAlaThrThrThrThrProGlyTh 268
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
713 .....CAGGAGAAATTTGCGG 728
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
268 rSerThrSerProLysPheHisThrThrThrThrThrThrThrGlnAla 285
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
729 CCGAGCGGATCGGCTGACGAGGTAAAGCGCAAGCTTCTGCTGTTA 778
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
285 LHisPheLysProCysArg.....AspLysAspLeuAlaArg 297
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
779 TGC...ACGGCTTGGGTCTGCTTCCACCGAAACA..... 811
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
298 CysLeuAsnAspGlyLeuCysPheValIleGluThrLeuThrGlySerH 314
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
812 ...AGATGGCGCGCATCAAGATTTGGCAGATATGGCGCAACATCAAGA 857
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
314 sArgHisCysArgCysLysGlnGlyThrGlnGlyValAlaArgCysAspGln 331

```

```

858 CTATGCCCGCAGACGACCATCCCGGATTTGGGAGATCCAAACCCCAATGCC 907
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
331 heLeuProLysThrAspSerIleLeuSerAsp.....ProThrAsp 344
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
908 CAC..... 910
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
345 HisLeuGlyIleGluPheMetGluSerGlnGluValArgGlnAla 361
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
910 ..... 910
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
361 IleuSerIleSerCysIleIlePheGlyIleValIleValGlyMetPheC 378
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
911 .....AAGCATAGAAAGCCCTCAGCAATA 934
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
378 ySalAlaPheThrPheLysSerLysGlnAlaLysGlnIleGlnGlu 394
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
935 TCTTTACGCGAGTCATCCCGCTCAAGAGGATTTGAGCTGTGGGGAATA 984
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
395 GlnLeuLysValProGlnAsnGlyLysSerLysSerLeuLysAlaSer 411
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
985 TACGGCTTGGGCGGATCAGCGCACATCTGTCAAGCGGTCGAGATGG 1034
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
411 rThr.....MetAlaLysSerIleLysLeu 420
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1035 CGAGATCGCATTTGCCGAAAGGGAAT.....CCGCCG 1066
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
420 alyLysSerHisValGlnLeuGlnAsnThrLysLysValGluArgHisPro 436
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1067 TCAGCGACATTTTCCGATGCGGCATACGCGCAAAATCCGCTCCCTTAC 1116
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
437 ValThrAlaLeuGlnLysMetMetGluSerSerPheValGlyProGln 453
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1117 CATTCGCGCAATATTCGTTCAACTTGAGCGAGGCTTACGCGCAAGAAA 1166
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
453 rPheProGluValProSerProAspArgGlySerGlnSerValLysHis 470
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1167 CATCACCCTCTCAACCGCTGCCCGCTCAAGCGCAAAAGATGTGAAATCG 1216
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
470 lsArgSerLeuSerSerCysLysSerProGlnArgSerGlyMetLeu 486
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1217 CAACAACAAGCCACCCGAGCA.....CGAAAGTCCGCTTG. 1252
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
487 HisArgAsnAlaPheArgArgThrProProSerProArgSerArgLeu 503
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1253 .ACGGTAAAGGTTCCGAATTTGAAAAGAGCTAAATATCATAGACG 1301
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
503 yGlyIleValGly..... 507
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1302 AATTAAATACCGCTGTACACAGTGAATCCTATAGATGAACCGCTTTTA 1351
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
508 .....ProAlaThrGlnGln..... 512
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1352 ATCTAAAGGTTCTGTGGATCGGCTATTTGGTTATTAACGCCACA 1401
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
513 .....LeuGlnLysSerArgIle.....ProAs 520
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1402 ATTCAATACGCAAAATATTCACAGGC...AAGGTAGAAATCATATATCC 1448
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
520 pGlnAspThrIleProCysGlnGlyIleGluValArgLysThrIleSerH 537
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1449 ACSTAAAAATTTACTCTC.....CTTACGACCGCTAC 1480
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
537 lsLeuProIleGlnLeuThrProCysValGlnArgProLeuAspLeuLys 553
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1481 CAAGAAGCACTAATATGATGATTTGGATTAATTTGGTAATGATGACT 1530
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
554 SerSer.....SerLys 558
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
1531 AAG.....GTCCATCAAGAACTAAAGTCAAAAT 1562
    |||:|  |||:|  |||:|  |||:|  |||:|  |||:|
558 ulysThrGlnArgAsnThrSerIleAsnMetGlnLeuProSerArgGlu 575

```

1563 TGAATGGAGTTCATTTCTATAAAGAGAGACCACTTGATGG 1612
 575 hrAsnProTyrPheAsnSerLeuGlulnysasp..... 586
 1613 CTAGTAGCGATGGTAAC 1630
 587 LeuValIleTyrSerSer 592

seq_name: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:us-08-287-001A-2

seq_documentation_block:

; Sequence 2, Application US/08287001A
 ; Patent No. 5622861
 ; GENERAL INFORMATION:
 ; APPLICANT: KAPLAN, GERARDO
 ; APPLICANT: FEINSTONE, STEPHEN M.
 ; TITLE OF INVENTION: HEPATITIS A VIRUS RECEPTOR AND METHODS
 ; TITLE OF INVENTION: OF USE
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
 ; STREET: Suite 1200, The Candler Bldg, 127 Peachtree
 ; STREET: Street
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: USA
 ; ZIP: 30303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/287,001A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Spratt, Gwedolyn D.
 ; REGISTRATION NUMBER: 36,016
 ; REFERENCE/DOCKET NUMBER: 1414,621
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 404/688-0770
 ; TELEFAX: 404/688-9880
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 451 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-287-001A-2

alignment_scores:

Quality: 112.50 Length: 255
 Ratio: 1.023 Gaps: 13
 Percent Similarity: 43.137 Percent Identity: 23.922

alignment_block:

US-09-303-518D-465 x US-08-287-001A-2 ..

Align seg 1/1 to: US-08-287-001A-2 from: 1 to: 451

13 CGCAAAATATCCCTATTCGTGCATACGCGAGTG..... 48
 86 ArgAspValSerLeuThrIleAlaAsnThrAlaValSerAspSerGlyTyr 102
 49TGCCTGGCGCATGCATGCACAG..... 70
 102 eTyrCys.CysArgValLysHisSerGlyTyrPheAsnAspMetLysIle 118
 71CCTCAGATTGGCAACAGTCTTT 96
 119 ThrIleSerLeuLysIleGlyProPheAlaValThrThrProIleValArg 135

97 ATCCGCGAGGTTCTCGACCGTCAGCATTTGCAGACCGAGGAAATACCA 146
 135 gThrValAlaGlyThrSerThrThrValProThrThrThrLeuProThr 152
 147 CCTATTGCGCAGCAGGAGGGAACCTTGCCAGCGCAGCGGTATATCGAT 196
 152 hrThrThr.....LeuProThrThrThrThrLeuProThr 163
 157 TGGGAACATACAAAGCCATCAGTTGGCAACCTGTTTCATCCAGACGCG 246
 164 ThrThrThrLeuPro.....MetThrThrThr 173
 247 GCCATTAAAGAAATATGCGTACATTCGCGCTTTCCGATCAGCGGCA 296
 173 uProMetThrThrThrLeuProThrThrThrValProThrThrThr 190
 297 CGAAGTCCATTCCTCCCTCGACAAACCATGCTTCATTCGATTCGTGATG 346
 190 hr.....LeuProThrThrThrThrPro..... 198
 347 AAGCCGTAATCCGCTGACGATTCAGCTTACCCGATTCATTCGAGAC 396
 199 ...ThrThrLeuProMetThrThrThrLeuProThr..... 209
 397 GGATACGAAACACCATCCG.....CCGACGCTATGACCG 431
 210 ...ThrThrThrLeuProThrThrThrLeuProThrThrMetThr. 224
 432 GCCACAGGCGCGGCTATCCCGCTCCCAAGCGCGAGATATATACA 481
 225LeuProMetThrThrThrLeuProThr 233
 482 GCTACGACATAAAGCGCTTGCCCAAAATATCCGCTACCTACCGAC 531
 234 ThrThrThr.....LeuPro.....ThrThrThrThrLeuProThr 245
 532 AACCCGACCGCGACAGACGCGTGTGCGAGCGTATCCCATATACCGGAG 581
 245 ThrThrLeuProThrMetThrThrThrThrThrThrThrThrThr 262
 582 TATGCTGACGAGAGTAGCGACGATTCAAACGCGCCACCGATACA 631
 262 eMet.....Thr 264
 632 GCCCGACGTGACAGATCGGGCAATGCCCGCAAGCTTTACAGCGCACT 681
 265 LeuProThrThrThrThrLeuProThrThrThrThrLeuProThrThr 281
 682 GCAGATATCGTCA 694
 281 MetValSerThr 285

seq_name: /cgn2_6/ptodata/1/1aa/PCBUS_COMB.pep:PCB-US95-09941-2

seq_documentation_block:

; Sequence 2, Application PC/TUS9509941
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: HEPATITIS A VIRUS RECEPTOR AND METHODS
 ; TITLE OF INVENTION: OF USE
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
 ; STREET: Suite 1200, The Candler Bldg, 127 Peachtree
 ; STREET: Street
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: USA
 ; ZIP: 30303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09941
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/287,001
; FILING DATE: 5 AUG 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwedolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-09941-2

```

```

alignment_scores:
  Quality: 112.50      Length: 255      Gaps: 13
  Ratio: 1.023
  Percent Similarity: 43.137      Percent Identity: 23.922

```

```

alignment_block:
  US-09-303-518D-465 x PCT-US95-09941-2 ..

```

```

Align seg 1/1 to: PCT-US95-09941-2 from: 1 to: 451

```

```

13 CGCAAAATATCCCTTATCTGTCATCTGGCAGT..... 48
14  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
86 ArgspValSerLeuThrIleAlaAsnThrAlaValSeraspSergIyl 102
49  ....TGCGCCGATGCATGCACAG..... 70
102 eTfYcys.CysArgValIlyshHisserIylTrpHeasnaSpMetIyle 118
71  ....CCTCAGATTGGCAACGATTCCTTT 96
119 ThrIleSerLeuIlyIleGlyProProArgValThrThrProIleValar 135
97 ATCCGGCAGGTTCTCGACCGTCAGCATTTGCACCCGACGGAANAACA 146
135 gThrValArgThrSerThrThrValProThrThrThrLeuProThrThr 152
147 CCTTTGCGGACGACGGGGGAACCTTGGCAGGCGGCGTCATATCGGAT 196
152  |||  |||  |||  |||  |||  |||  |||  |||  |||
197 TGGGAACATACAAAGCATCATGTTGGGCAACCTGTTCATCAGACGCG 246
164 ThrThrThrLeuPro.....MethrThrThrIle 173
247 GCCATTAAGGAATATCGCTACATTCGCTTTCCGATCAGGGCA 296
173 uProMethrThrThrLeuProThrThrThrThrValProThrThrThr 190
297 CGAGTGTCATTCGCTTCGACACACATGCTTCATTCGATTCATG 346
190  |||  |||  |||  |||  |||  |||  |||  |||  |||
347 AAGCGGTAGTCCGTTGACGATTCAGCTTTACCGCATCATTCGGAC 396
199  ....ThrThrLeuProMethrThrThrLeuProThr..... 209
397 GGATACGACACCATCCG.....CGACGCGTATGACGG 431
|||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

210  ....ThrArgThrLeuProThrThrThrThrLeuProThrThrMethr. 224
432 GCCACAGGGGCGGCGCTATCCCGCTCCCAAGCGCAGGATATATACA 481
225  ....LeuProMethrThrThrThrLeuProThr 233
482 GCTACGACATAAAGCGTTGCCCAAAATATCCGCTTACCTGACGAC 531
234 ThrThrThr.....LeuPro.....ThrThrThrThrLeuProTh 245
532 AACCGACGACCGGACAAAGCGTTCGACGCTTCCCAATATCCGAG 581
245 rThrThrLeuProThrMethrThrLeuProThrThrThrThrLeuPro 262
582 TATGCTGACGCAAGAGTAGGCGACGATTCAAACGCCACCCGATACA 631
262 eMet.....Thr 264
632 GCCCGACGCTGACAGATCGGCAATGCCCGCAAGCTTTCACGCGACT 681
265 LeuProThrThrThrThrLeuProThrThrThrThrLeuProThrThr 281
682 GCAGATATCGTCA 694
281 rMetValSerThr 285

```

```

seq_name: /cgn2_6/ptodata/1/laa/6A_COMB.pep:US-09-007-999-2

```

```

seq_documentation_block:
; Sequence 2, Application US/0907999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007,999
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-007-999-2

```

```

alignment_scores:
  Quality: 112.00      Length: 610
  Ratio: 0.424
  Percent Similarity: 43.279      Percent Identity: 20.984

```

```

alignment_block:
  US-09-303-518D-465 x US-09-007-999-2 ..

```

```

Align seg 1/1 to: US-09-007-999-2 from: 1 to: 1475

```

```

64 GCACACGCTCAGATTGGCAACGATTCCTTTATCCGCGAGTCTCGA 113
617 SerTyrAlaIleuLeuThrAsnIlySerSerValProArgValTyrTy 633
114 CGTGCGATTCGCAACCGACGGGGAATAC..... 144
633 rIlyaspMetPheThrAspAspIlyGlnTyrMetAlaHisIlyThrIle 650
145  ....CACCTATTC 153
650 snTyrGluAlaIleGluThrLeuLeuIlyAlaArgIleIlySerValSer 666
154 GGCAGCAGGGGGAACCTTCCGACGCGACGCGTCA..... 189
|||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

667 GlyGlyGlnAlaMetArgAsnGlnGlnValGlyAsnSerGluIleIleThr 683
190ATGGATTTGGGAAACATACAAAGCCATCAGTTGGGGAACC 229
683 rSerValArgTyrGlyLysGlyAlaLeuLysAlaThrAspThrGlyAspA 700
230 TGTTCATCCGACAG.....GGCGCATTTAAAGAAATATCGGCTAC 270
700 rGthrThrArgThrSerGlyValAlaValIleGluLysAsnAspProSer 716
271 ATT..GTCCGCTTTCCGAT.....CA 290
717 LeuAlaLeuLysAlaSerAspArgValValAlaAsnMetGlyAlaAlaThr 733
291 CGGGACAGAGTCCATTCCTCCCTTC.....GACAC.....C 322
733 sLysAsnGlnAlaIleArgProLeuLeuLeuThrThrAspAsnGlyIleI 750
323 ATGCTCACATTCGATTCGATGAGCGCGGTAGTCCCGTTGACGGATTTC 372
750 ySalATyrHisSerAspGlnGlnAlaAlaGly..... 760
373 AGCTTTACCGCATCCATTGGGACGATACGAA.....CACCATCCGC 416
761 ...LeuValArgTyrThrAsnAspArgGlyGluLeuIlePheThrAlaAl 776
417 CGAG.....GGCTATGACGGCCACAGGGCGGCTATCC..... 453
776 AspIleLysGlyTyrAlaAsnProGlnValSerGlyTyrLeuGlyValT 793
454 ..GCTCCAAAGGCGGAGGATATATACAGCTACGATCAAAAGGCGTT 501
793 rPalProValGlyAlaAlaLeuIleLysMetPheAlaLeuArg..... 807
502 GCCCAAAATATCCGCTCAACCTGACCGACACCGACCGGACACAG 551
808LeuAlaArgProHisG 813
552 GCTTTGACCGCTTCCACAATACCGGTAGTATGACGACGAGGATG 601
813 ngInMetAlaSerValHisGlnAsnAlaAlaLeuAspSerArgValMet 830
602 GCGACGATTCAAAGCGCCGACCGATACAGCCCGAGCTGACAGATCG 651
830 heGluGlyPheSerAsnPheGlnAlaPheAlaThrLysLysGluGlyTyr 846
652 GCGAATGCCCGACGCTTCAAC.....GCGAC 680
847 ThrAsnValValIleAlaLysAsnValAspLysPheAlaGluTrpIleVal 863
681 TGCAGATATGTCAAAACATCATCGCGCGGAGGAGAAATGTGCGGG 730
863 lThrAsp.....PheGluMetAlaProGlnTyrValSer 875
731 CAGCGCATGCCGTG.....CAGGATATAGCAGAGCTCAACAT 771
875 erThrAspGlySerPheLeuAspSerValIleGlnAsnGlyTyrAlaPhe 891
772 GCTGTATGACGCGTTGGCTGCTTTCACCGCAAAACAAGTGGCGG 821
892 ThrAspArgTyrAsnLeuGly.....IleSerLysProAsnSerTyrIle 907
822 CATCAACGATTTGGCAGAT...ATGGCGCAATCAAGACTATGCGCGAG 868
907 rAlaAspAspLeuValLysAlaIleLysAlaLeuHisSerLysGlyIle 924
869 CAGCCATCGCGATTTGGGACGTCCAAAACCCCAATGCCGACAGGACAT 918
924 yValMetAlaAspTrpValProAspGlnMetTyrAlaPheProGluLys 940
919 GAAGCCCTGACCAATATCTTACGGCAGTCATCCCGTCAAGGAGATTGG 968
941 GluValValThr..... 944

969 AGCTGTTCG...GGAATACGGCTTGGGGGACATCAGCGACATCTG 1015
945 ..AlaThrArgValAspLysTyrGlyThr.....Prov 955
1016 TCMAAGGCTCGCAGATG...GGCGAGATCGCATTCGCGAAGGAAATCC 1062
955 AlAlaGlySerGlnIleLysAsnThrLeuTyrValValAspGlyLysSer 971
1063 GCCGTACCGACAT.....TTGCCGATCGGC 1091
972 SerGlyLysAspGlnGlnAlaLysTyrGlyAlaPheLeuGluGluLe 988
1092 ATACGCCAATACCGTCCCTTACCATTCGCGAATTCGTTCAAACT 1141
988 uGlnAlaLysTyrProGluLeuPheAlaArgLysGlnIle..... 1001
1142 TGGACACCGTTACGGCAAGAAACATCACCTCCACACCGTCCG... 1188
1002SerThrGlyValPromet 1007
1189 ..CCGTC.....AACGAAA 1202
1008 AspProSerValLysIleLysGlnTrpSerAlaLysTyrPheAsnGlyTh 1024
1203 GAATGTAAACTGGCAACAAACGCCACCGACCAACCAAGTGGCGTTG 1252
1024 rAsnIleLeu..... 1027
1253 ACGGTAAAGGTTTCCGATTTTGAAGAAAGCGTAATAATACGATACGAG 1302
1028 ..GlyArgGlyAlaGlyTyrValLeuLysAspGlnAlaThrAsnThrTyr 1043
1303 ATTATACCGCT...GTACCAAGTGAATCTATAGTGAACCGCTT 1349
1044 PheAsnIleSerAspAsnLysGlnIleAsnPheLeuProLysThrLeuLe 1060
1350 TATCTTAAAGGTTCTGTCGATCGCTCATTTGTCGTAACTGCA 1399
1060 uAsnGlnAspSerGlnValGlyPheSerTyrAspGlyLysGlyTyrVal 1076
1400 GAATTCATACGCAAAATTACCAAGCAAGGTAGATCAGATATATCCCA 1449
1077 TyrTyrSerThrSerGlyTyrGlnAlaLys...AsnThrPheIleSerG 1092
1450 CCTAAAAATTACTCTCTTCAGACCGCTACCAAAAGACCTAATAAG 1499
1092 u.....GlyA 1094
1500 ATATTGGATAATTGGTAATGAATGAGCTAAAGTCCATCAAGAACTA 1549
1094 sPlyTrpTyrTyrPheAspAsnAsnGlyTyrMetValThrGlyAlaGln 1110
1550 AAGGTCAAGAAATTTGAATGGATGTTG.....AATTGCTAANAACAGA 1593
1111 SerIleAsnGlyValAsnTyrTyrPheLeuSerAsnGlyLeuGlnLeuAr 1127
1594 AGAGAC.....AAGTGGATGG 1612
1127 gAspAlaIleLeuLysAsnGluAspGly 1136
seq_name: /cgn2_6/ptodata/1/laa/6b_COMB.pep:us-09-210-361-2
seq_documentation_block:
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11

```
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRF
; ORGANISM: Streptococcus mutans
US-09-210-361-2
```

```
alignment_scores:
  Quality: 112.00      Length: 610
  Ratio: 0.424        Gaps: 32
  Percent Similarity: 43.279  Percent Identity: 20.984
```

```
alignment_block:
US-09-303-518D-465 x US-09-210-361-2 ..
```

```
Align seg 1/1 to: US-09-210-361-2 from: 1 to: 1475
```

```
64 GCACACCCCTCAGATTGGCAACAGATTCTTTATCCGCGAGTTCGCA 113
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
617 SerTyrAlaLeuLeuLeuThrAsnLysSerValProAlaGlyTyr 633
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 CCGTCACGATTTGCAACCCGAGGAATAC.....144
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
633 rGlyAspMetPheThrAspArgLysGlyMetAlaHisLysThrIleA 650
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
145 .....CAACATATTC 153
650 snTyrGluAlaIleGluThrLeuLysAlaArgIleLysTyrValSer 666
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
154 GGCAGCAGGGGGGAATCTGCCAGCGCGCAT.....189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
667 GlyLysGlnAlaMetArgAsnGlnGlnValGlyAsnSerGluIleLeu 683
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
190 .....ATCGGATTGGGAACATACAAAGCCATCAGTTGGCGCAAC 229
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
683 rSerValArgTyrGlyLysGlyAlaLeuLysAlaThrAspThrGlyAsp 700
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
230 TGTTCATCCAGCAG.....GGCGCATTAAGGAATATCGGGTAC 270
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
700 rGlnThrAlaGlnThrSerGlyValAlaValIleGluGlyAsnAsnPro 716
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
271 ATT...GTCCGCTTTCCGAT.....CA 290
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
717 LeuArgLeuLysAlaSerAspArgValValAlaAsnMetGlyAlaHis 733
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
291 CGGGCAGCAAGTCCATCCCCCTTC.....GACAAAC.....C 322
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
733 sLysAsnGlnAlaLysArgProLeuLeuLeuThrThrAspAsnGlyLeu 750
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
323 ATGGCTCACAATTCGATTGTGATGAACCGGTAGTCCCGTTGACGATTC 372
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
750 ySalAlaTyrHisSerAspGlnGlnAlaIleGly.....760
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
373 AGCCATTACGCAATTCATTTGGACGGATAGCA.....CACCATCCGC 416
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
761 ...LeuValArgTyrThrAsnAspArgGlyGluLeuIlePheThrAla 776
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
417 CGAC.....GGCTATGACGGGCCACAGGGCGCGGCTATCC..... 453
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
776 aspIleLysGlyTyrAlaAsnProGlnValSerGlyTyrLeuGlyAla 793
454 ..GCTCCCAAGCGCGAGGATATATACGCTACAGATTAAGCGCTT 501
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
793 rValProValGlyAlaAlaLeuIleLysMetPheAlaLeuTyr..... 807
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
502 GCCCAAAATATTCGCTCAACCTGACGCAACACCGCAGCGCAACG 551
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
808 .....LeuAlaArgProHisGlu 813
552 GCTTGTCCAGCGTTTCCACATACCGGTAGTGTGACGCAAGAGTAG 601
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
813 nGlnMetAlaSerValHisGlnAsnAlaAlaLeuAspSerArgValMet 830
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
602 GCGAGGATTCAAACGGCGCACCGGATACGCGCCGAGCTGGACGATCG 651
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
830 heGluGlyPheSerAsnProGlnAlaPheAlaThrLysLysGluLeuTyr 846
652 GGCATGCGCGCCGAGCTTTCAC.....GGCAC 680
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
847 ThrAsnValValIleAlaLysAsnValAspLysPheAlaGluTrpGlyVa 863
681 TGCAGATATCGTCAAAAACATCATCGCGCGGACGAGAAATGTGCGCG 730
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
863 lThrAsp.....PheGluMetAlaProGlnTyrValSerS 875
731 CAGGGATGCGCGTG.....CAGGGATTAAGGAGCGTCAACAT 771
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
875 erThrAspLysPheLeuAspSerValIleGlnAsnGlyTyrAlaPhe 891
772 GCTGTATGACGCGCTGTGGTCTGCTTCCACCGAAACAGATGGCGCG 821
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
892 ThrAspArgTyrAspLeuGly...IleSerLysProAsnLysTyrGlyTh 907
822 CATCAACGATTTGGCAGAT...ATGGCGCACTCAAAAGCTATGCGCGAG 868
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
907 rAlaAspAspLeuValLysAlaIleLysAlaLeuHisSerLysGlyLeu 924
869 CAGCCATCCGCGATTTGGCGATGCAAAACCCCAACGCGSCCAAGCAT 918
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
924 yValMetAlaAspTrpValProAspGlnMetTyrAlaPheProGlnLys 940
919 GAAGCCGTACGCAATATCTTACGGCAGTCAATCCCGTCAAGGGATTGG 968
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
941 GluValAlaThr.....944
969 AGCTGTGGG...GGAAATACGCTTGGCGCGCATCAGGCACATCCTG 1015
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
945 .AlaThrArgValAspLysTyrGlyThr.....Prov 955
1016 TCACGCGGTGCGCAGATG...GGCGAGATCGCATTCGGAAGGAAATCC 1062
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
955 AlaIleGlySerGlnIleLysAsnThrLeuTyrValValAspLysSer 971
1063 GCCGTACGCGCAAT.....TTGGCGATCGCGC 1091
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
972 SerGlyLysAspGlnGlnAlaLysTyrGlyAlaPheLeuGluGluLe 988
1092 ATACGCCAAATACCGCTCCCTTACCATTCGCGGAATTCGGTTCAACT 1141
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
988 nGlnAlaLysTyrProGlnLeuPheAlaArgLysGlnIle..... 1001
1142 TGGACGAGGTACGCAAGAAAGAAACATCACCTCCACACGTCGCG... 1188
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1002 .....SerThrGlyAlaProMet 1007
1189 ...CCGTCA.....AAGCGAA 1202
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1008 AspProSerValLysIleLysGlnTrpSerAlaLysTyrPheAsnGlyTh 1024
1203 GAATGTGAACATGCGCAACAAAGCGCACCGGACCAAGCAAAAGTCCG 1252
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1024 rAsnIleLeu.....1027
```



```

1253 ACGTAAGGTTCCGAATTTTGAAAGACGTAATAATACATACGAGA 1302
      |||||
1028 ..GlyArgGlyAlaGlyTyrValLeuLysAspGlnAlaThrAsnTrpTyr 1043
      |||||
1303 ATTAATACCCCT...GTACCACAAGTGAATCTATAGATGACCCGTGT 1349
      |||||
1044 PheAsnIleSerAspAsnLysGlnIleAsnPheLeuProLysThrLeu 1060
      |||||
1350 TAATCCTAAGGTTCTGCGATCGGCTCATCTTGTCTATACAGCCA 1399
      |||||
1060 uasnGlnAspSerGlnValIlePheSerTyrAsp..GlyLysGlyTyrVal 1076
      |||||
1400 GAATCAATACGCAAAATTAACGAAGCAAGTGAATCAGATATATCCA 1449
      |||||
1077 TyrTyrSerThrSerGlyTyrGlnAlaLys..AsnThrPheIleSerG 1092
      |||||
1450 CCAAAATTAATCTCTCTTCAGACCGCTACCAAAAGACCTAATATG 1499
      |||||
1092 u.....GlyA 1094
      |||||
1500 ATATTGATAAATTGTGATGATGACTAAGCTCAAGTCCATCAAGACTA 1549
      |||||
1094 splYstrpTyrTyrPheAspAsnAsnGlyTyrMetValThrGlyAlaGln 1110
      |||||
1550 AAGCTCAAGATTTGAATGGGATGTTT.....AATTGCTAAACAGCA 1593
      |||||
1111 SerIleAsnGlyValAsnTyrTyrPheLeuSerAsnGlyLeuGlnLeu 1127
      |||||
1594 AGAGAGC.....ACTTGATGGG 1612
      |||||
1127 gasPaLaIleLeuLysAsnGlnAspGly 1136

```

seq_name: /sgn2_6/ptodata/1/1aa/5B_COMB.pep:US-08-224-482-4

seq_documentation_block:

```

: Sequence 4, Application US/08224482
: Patent No. 5837692
: GENERAL INFORMATION:
: APPLICANT: Mercola, Dan
: APPLICANT: Adamson, Eileen D.
: TITLE OF INVENTION: Inhibition of the Mitogenic Activity of
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Campbell and Flores
: STREET: 4370 La Jolla Village Drive, Suite 700
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92122
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/224,482
: FILING DATE: 07-APR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Campbell, Cathryn A.
: REGISTRATION NUMBER: 31,815
: REFERENCE/DOCKET NUMBER: P-ME 9913
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-9001
: TELEFAX: (619) 535-8949
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 543 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
US-08-224-482-4

```

alignment_scores:

Quality:	111.50	Length:	430
Ratio:	0.616	Gaps:	19
Percent Similarity:	42.093	Percent Identity:	21.163

alignment_block:

US-09-303-518d-465 x US-08-224-482-4 ..

Align seg 1/1 to: US-08-224-482-4 from: 1 to: 543

```

77 ATTTGCAAGCATCTTTATCCGCGAGTTCTCGACCGTCACATATTC 126
      |||||
149 LeuTrpProGluProLeuPheSerLeuValSerGlyLeuValSerMetThr 165
      |||||
127 GAACCCGACGGGAATACCACTATTGCGACGACGAGGGGGAGACTGCGCA 176
      |||||
165 rAsnPro.....ProLaserSerSerSerAlaProS 176
      |||||
177 GCGCAGCGGTCATATCGGATTGGGAACATACAAAGCATCAGTTGGGCA 226
      |||||
176 ePrAlaAla.....Ser 180
      |||||
227 ACCTGTCATCCAGCAGCGGCGCATTAAGGAATATCGCTACATTTGTC 276
      |||||
181 SerAlaSerAlaSerGlnSerPro.....ProLeuSerCysAlaValPro 195
      |||||
277 CGCTTTTCCGATCACGGGCACGAGAGTCCATTCCTCCCTTCGACAAACATGC 326
      |||||
195 oSerAsnAspSerSerProIleTyrSerAlaAlaProThrPheProThrP 212
      |||||
327 CTCACATTCGATTCGTCTGATGAAGCGGAGTACGCCGTTGACGATTCAGCC 376
      |||||
212 roAsnThrAspIlePheProGluPro.....GlnSerGlnAla 224
      |||||
377 TTTACCGCATCCATTGGGACGAGATACGAACACATCCCG.....CC 417
      |||||
225 PheProGlySerAlaGlyThrAlaLeuGlnTyrProProAlaTyrPr 241
      |||||
418 GACGCTTATGACGGGCCACAGGCGGCGGCTATCCG.....CTCC 458
      |||||
241 oAlaIleLysGlyGlyPheGlnValProMetIleProAspTyrLeuPheP 258
      |||||
459 CAAGCGCGAGGATATATA..... 479
      |||||
258 ro.GlnGlnGlnGlyAspLeuGlyLeuGlyThrProAspGlnLysProPh 274
      |||||
479 ..... 479
      |||||
274 eGlnGlyLeuGlnLysSerArgThrGlnLysProSerLeuThrProLeuSer 291
      |||||
479 ..... 479
      |||||
291 hTrIleYsAlaPheAlaThrGlnSerGlySerGlnAspLeuYsAlaLeu 307
      |||||
480 .....CAGCTACGACATATAAAGGGGTTGCCCAATA 511
      |||||
308 AsnThrSerTyrGlnSerGlnLeuIleLysProSerThrGlnMetAlaGlySty 324
      |||||
512 TCCGCTCAACCTGACGACGACGACGACGACGACGACGACGACGCTTGTCAC 561
      |||||
324 rProAsnArgProSerLysThrProProHISGln.....ArgP 337
      |||||
562 CGTTTCCCAATACCGGATGATGCTGACGACG...AGTAGGCGACGG 608
      |||||
337 roTyrAlaCysProValGlnSerCysAspArgArgPheSerArgSerAsp 353
      |||||
609 ATTCAAAGCGGCACCGCATACAGCCCGGAGCTGACAGATCGGGCATG 658
      |||||
354 GluLeuThrArgHis...IleArgIleHisThrGlyGln..... 365
      |||||
659 CCGCGAAGCTTTCAACGCGACTGACAGATATCGTCAAAACATCATCGGC 708

```

```

366 ...LyPProheIn.....CysArgIleCysMetArgAspPheSerA 379
709 GCGGCGAGAGAAATTTGCGGCGAGCGGATCCGTCAGGGATTAAGCGA 758
379 rGserAspHisIleuThrThiHisIleArgThrHisThrGlyLysPro 395
759 AGGTCGAACATTCGTTGATGACGGCTGGGTGCTTCCACCGGAA 808
396 PheAlaCysAspIleCysGlyArgLysPheAlaArgSerAspIleuArg 412
809 ACAAGATGCGCGCATCAACGATTTGGCAGATAGCGGCACTCAAGAC 858
412 s.ArgHisThrLysIleHis.....LeuArgGlnLysAsp 423
859 TATCCCGCAGACGATCCGCGATGGCGATGCCAAACCCCATGCCGC 908
424 LysLysAla.....AspLysSerValValAlaLaserSerAlaTh 436
909 ACAAGGATAGAACGCGTCAGCATATCTTACGGCAGTCATCCCGTCA 958
436 rSerSerLeuSerSerTyProSerProValAlaIleThrSerTyPro... 451
959 AAGGATTTGGAGCTTTGGGGAATATCGGCTTGGCGCATCACGGCA 1008
452 .....SerProValThrThrSerTyProSerProAlaIleThrSer 465
1009 CATCTGTCAAGCGGTGCGAGATGGCGAGATGCGCATGCCGAAGGGA 1058
466 TyrPro..... 467
1059 ATCCGCGCTGAGCGACATTTTGGCGATGGCGATACGCCAATACCGGT 1108
468 .SerProValProThrSerPheSerSerProGlySerSerThrTyPro 484
1109 CCCCTTACATTCGCCGAATATCCGTTCAACTTGGAGCAGCGTTACGGC 1158
484 erProValHisSerGlyPheProSerProSerValAlaIleThrTy... 499
1159 AAGAAGAAACATCACCTCTCTCAACCGTGGCGCGCGCA 1194
500 .....SerSerValProProAla 505

seq_name: /cgn2_6/plodata/1/1aa/5B.COMB.pep:US-08-557-139-2
seq_documentation_block:
; Sequence 2, Application US/08557139
; Patent No. 5827730
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Oluf
; APPLICANT: Bjorbak, Christian
; APPLICANT: Frederiksen, Kathrine A.
; TITLE OF INVENTION: MUTANT DNA ENCODING INSULIN RECEPTOR
; TITLE OF INVENTION: SUBSTRATE 1
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 58277300 No. 5827730disk of No. 5827730th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,139
; FILING DATE: 12-FEB-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.

```

```

REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4041.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 867-0123
TELEFAX: (212) 878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1243 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-557-139-2

alignment_scores:
Quality: 111.50 Length: 424
Ratio: 0.603 Gaps: 25
Percent Similarity: 43.632 Percent Identity: 23.349

Alignment block:
US-09-303-518D-465 x US-08-557-139-2 ..
Align seg 1/1 to: US-08-557-139-2 from: 1 to: 1243

105 GGTCTCGACGCTGACATTTGGAACCGGAGGAATATGACCTATTGCG 154
319 GlyLysLysProGlySerPheArgValAlaArgLaser..... 330
155 GACAGAGGGGGAAGCTTGGCGAGCGGACGCGATATGCGATTGGGAAC 204
331 .....SerAspGlyLysLysLys 337
205 ATACGAAGCCATGAGT.....GGCAACCTGTTCATCA..... 239
337 eSerArgProLaserValAspLysSerProValSerProSerThrAsn 353
240 .....GCAGCGGCGCATTAAGAAATATGCGCTACATGTGCGCT 280
354 ArgThrHisAlaIleAlaArgHisArgGlyAlaArgLysLeuHisProPole 370
281 TTTCGATGACGG.....G 294
370 uAsnHisSerArgSerIleProMetProAlaSerArgCysSerArgSerA 387
295 CACGAAGTCCATCCGCTTCGCAAC.....CATGCTC 329
387 lathSerProValSerLeu.SerSerSerSerThrSerGlyHisGlySe 403
330 ACATTCCGAT.....TCTGATGAAGCGGTA 355
403 rThrSerAspCysLeuPheProArgArgSerSerSerValSerGlyS 420
356 GTCCGCTGACGATTCAGCTTACCGCATTCATGGACGAGCAAGCAA 405
420 erProSerAspGlyLysPhe.....LaserSerAspGlyLys 433
406 CACATTCGCGCGGAGCTATGACGG.....CCACAGG 440
434 SerSerProCysAspPheArgSerSerPheArgSerValThrProAsp 450
441 CGGCGCTAT...CCGCTCCCAAGCGGAGGATATATACASTAG 487
450 rLeuGlyHisThrProProAlaArgGlyLysLysLysLeuSerAsnTy 467
488 ACATTAAGGCGTGGCCCAAAATATCCGCTGACGCTGACGAGCAAGCG 537
467 lCysMetGlyLysGlyProSerThrLeuThrAlaProAsn..... 481
538 AGCAGCGCAACAGGCTTGTGACGCGCTTTCACAAATACCGGTATGCT 587
482 .....GlyHisTyrlleuSerArgGlyLysLysLysLysLys 495
588 GACGCAAGAGTAGGCGAGGATTCAAACGCGCGCCGATGACGCGCG 637

```

```
495 sThProglYThrLeuLysly.....ThSerProA 506
      ||| |||:|||| |||
638 AGCTG.....GACGATCGGCAATCCCGCAGCTTTCACAGCGACT 681
      ||| |||:||||:|||||:|||||:|||||:
506 laLeuAlaGlyAspGluAlaAlaSerAlaLasp..LeuAspAsn.... 520
      ||| |||:||||:|||||:|||||:|||||:
682 GCAGATTCCTCAAAAACATCATCGCGC...CGGCAAGAGAAATTGTCG 728
      ||| |||:||||:|||||:|||||:|||||:
521 .ArgPheArgLysArgThrHisSerAlaGlyThrSerProThrLeuThr 537
      ||| |||:||||:|||||:|||||:|||||:
729 CGGAGGGGATGCGGTGCGGTATAGCGAAGCTCAACATTCGTGTTA 778
      ||| |||:||||:|||||:|||||:|||||:
537 tsGlnLysThrProSerGln..... 543
      ||| |||:||||:|||||:|||||:|||||:
779 TGCACGGCTGGGTGCTGCTTCCACGCAAAAAGATGGCCGCAATCAC 828
      ||| |||:||||:|||||:|||||:|||||:
544 .....SerSe 545
      ||| |||:||||:|||||:|||||:|||||:
829 GATTGGCAGATATGGCGCACTCAAGACTATGCCGACAGCCATCCG 878
      ||| |||:||||:|||||:|||||:|||||:
545 tValAlaSerIleGluGluThrGluMetMetProAlaTyProProG 562
      ||| |||:||||:|||||:|||||:|||||:
879 CGATTGGCGATGCCAAAACCCCAATGCCGACAAAGCATAGAACCGCTCA 928
      ||| |||:||||:|||||:|||||:|||||:
562 tYglYglSerGlyGlyArgLeuProGlyHis.....ArgHisSer 575
      ||| |||:||||:|||||:|||||:|||||:
929 GCATATCTTTACGCGCAGTCATCCCGTCAAAGAGATGGAGCTGTTCG 978
      ||| |||:||||:|||||:|||||:|||||:
576 AlaPheAlaProThrArgSerTYrProGluGlyLeuGluMet..... 590
      ||| |||:||||:|||||:|||||:|||||:
979 GGAATAATACGGCTTGGCGGCATACGCGCATCTGTCAAGCGGTGCA 1028
      ||| |||:||||:|||||:|||||:|||||:
591 .....HisProLeuGlu..... 594
      ||| |||:||||:|||||:|||||:|||||:
1029 GATGGCGGAGATGCATTCGCCGAAGGA.....AATCGCGCTCA 1069
      ||| |||:||||:|||||:|||||:|||||:
595 .....ArgArgGlyGlyHisHisArgProAspSer 604
      ||| |||:||||:|||||:|||||:|||||:
1070 GCGACAAAT.....TTGCCGATCGCGCATACGCCAAA 1101
      ||| |||:||||:|||||:|||||:|||||:
605 SerThrLeuHisThrAspAspGlyTYrMetProMet.....SerProG 619
      ||| |||:||||:|||||:|||||:|||||:
1102 TACCCGTCCTTACCATTCGCCGAATATCCGTTCAACTTGGAGCAGCG 1151
      ||| |||:||||:|||||:|||||:|||||:
619 yValAlaProValProSerGlyArgLysGlyAspTYrMetProMet 636
      ||| |||:||||:|||||:|||||:|||||:
1152 TTACGGGAAGAAACATCACTCCCTCA.....CCGTCGCCG 1189
      ||| |||:||||:|||||:|||||:|||||:
636 eSerProLysSerValSerAlaProGlnGlnIleIleAsnProIleArg 652
      ||| |||:||||:|||||:|||||:|||||:
1190 CGTCAACGGAAGAAATG 1207
      ||| |||:||||:|||||:|||||:|||||:
653 ArgHisProGlnArgVal 658
      ||| |||:||||:|||||:|||||:|||||:
seq_name: /cgn2.6/ptodata/1/iaa/5A.COMB.pep:US-08-325-267A-2
seq_documentation_block:
; Sequence 2, Application US/08325267A
; Patent No. 5585271
; GENERAL INFORMATION:
; APPLICANT: WAITARI, JUNJI
; APPLICANT: TAKAYA, YOSHIHIRO
; APPLICANT: OGAWA, MASAHIRO
; APPLICANT: PENTTILA, MERJA
; APPLICANT: ONNELA, MAJJA-LEENA
; APPLICANT: KERANEN, SIRKA
; TITLE OF INVENTION: YEAST AGGLUTINATION GENES AND YEAST
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400
```

```

; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,267A
; FILING DATE: 18-NOV-1994
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP PCT/JP94/00290
; FILING DATE: 24-FEB-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 38871/1993
; FILING DATE: 26-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2589-023-0XPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ. ID NO.: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1537 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-325-267A-2
```

```
alignment_scores:
  Quality: 111.50      Length: 399
  Ratio: 0.563         Gaps: 23
  Percent Similarity: 49.624  Percent Identity: 22.556
```

```
alignment_block:
US-09-303-518d-465 x US-08-325-267A-2 ..
Align seg 1/1 to: US-08-325-267A-2 from: 1 to: 1537
```

```
50 GCCTCCGATGCATGCACACGCTGACATTGGCAACGATCTTTATC 99
      ||| |||:||||:|||||:|||||:|||||:
498 AlaThrThrAlaMetThrThrThrGlnProTrp.....As 509
      ||| |||:||||:|||||:|||||:|||||:
100 CGGAGGTTTCGACCGTCAGCATTCGACCCGCGGAATACCACT 149
      ||| |||:||||:|||||:|||||:|||||:
509 nAspThrPheThrSerThrSer.....ThiGluMetThrTrp 522
      ||| |||:||||:|||||:|||||:|||||:
150 ATTGCGACAGAGGGGGGAGCTTCCGACGCGAGGTCATTCGATTG 199
      ||| |||:||||:|||||:|||||:|||||:
522 aThrGlyThrAsnGly..LeuProThrAspGluThrIleIleValIle 537
      ||| |||:||||:|||||:|||||:|||||:
200 GAACATATCAAAAGCCATCAGTTGGGCAACCTGTTTCATCCAGAGCGCC 249
      ||| |||:||||:|||||:|||||:|||||:
538 ArgThr.....ProThrThrAlaThrThrAlaMetThrThrGlnPr 552
      ||| |||:||||:|||||:|||||:|||||:
250 ATTAAGGAA.....ATATCGGCTACATTCGCGCTTTTCGATCAGG 293
      ||| |||:||||:|||||:|||||:|||||:
552 oTrpAsnAspThrPheThrSerThrSerThiGluMetThrThrValIle 569
      ||| |||:||||:|||||:|||||:|||||:
294 GCACGAGTCATTCCTCCCTTCGACAAACCATGCTCATATTCGATTG 343
      ||| |||:||||:|||||:|||||:|||||:
569 tYThrAsnGlyLeuProThrAspGluThr.....IleIleValIle 582
      ||| |||:||||:|||||:|||||:|||||:
344 ATGAAGCCGTAGTCCGCTGACG.....GA 369
      ||| |||:||||:|||||:|||||:|||||:
583 ArgThrProThrThrAlaThrThrAlaIleThrThrThrGluProTrpAs 599
```



```
LOCATION: (269)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (291)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (323)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (339)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (349)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (356)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (366)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (410)
OTHER INFORMATION: amino acid has not been identified
NAME/KEY: UNSURE
LOCATION: (421)
OTHER INFORMATION: amino acid has not been identified
US-09-156-836B-2
```

```
alignment_scores:
Quality: 116.50 Length: 343
Ratio: 0.932 Gaps: 19
Percent Similarity: 36.443 Percent Identity: 25.948
```

alignment_block:

US-09-303-518D-465/rev x US-09-156-836B-2 ..

Align seq 1/1 to: US-09-156-836B-2 from: 1 to: 430

```
888 TGCCCAATCCGCGATGCGTCGCGCATGATCTTTGCGCCATAT 839
||||| |||||
121 CysPro.....CysCysSer..... 125
838 CTGCCAATGTTGATGCGCGCATCTTTGCGTGAAGCAGAC.C 790
||||| ::|||::: ::|||:::
126 .CysGlnHis**gIyCysArgTyrCysArgTyrProGlySerArgTyrP 142
789 CAAGCCGTCGATACAGCAATGTTGCGCTGCTATACCGTCGACG 740
||||| ::|||::: |||||
142 roser.....SerArgCysProSerLeuArgCys..... 151
739 CATGCGCTGCGCGCAGCAATTTCTCCGCGCGCATGATGTTTTCAGC 690
||||| |||||
152 .....ArgArgPheArgCysProArg**ArgCys..... 161
689 ATATCTCAGTGCCTTGAAAGCTTCGCGCGCATGCCCATCTGTCCAG 640
||||| |||||
162 .....GlnArgTyrTyrCysProAsn**ThrG 171
639 CTGCGGGGTGATGCGGTGGCGCTTGAAATCCGCTGCTACCTCTTCG 590
||||| |||||
171 LysArgCys.....ArgCysProSerSer 179
589 TCA.....GCATACACCGGATGTTGAAAGCGTCGACAGCGCTGTGT 546
||||| ::||| |||||
180 SerArg**GlnTyr***SerProAlaGlyCysArgArgThrAlaArgCys 196
545 CCGGTGCTGCGGTGCGTCAAGTTGAGCGGATATTTTGGGCAAGCC 496
||||| |||||
196 sarGlyCysCysArgCys.....TrpArgT 206
495 TTTATGCGTAGCGTATATATCCCTCGCGCTTTGGAG..... 455
||||| |||||
```

```
206 hrArgCysCysCys.....ArgCysTrpGlnSerLeuGly 218
454 .....CGGATATGCGCGCGCCCTGCGC 432
219 **SerArgProArgSerArgSerArgArgCysSerArgArgPheG 235
431 CCGTATGCGCGTGGCGGATGTTGCTGATCCGT..... 395
235 nAsnArgCysCysArgSerArgGlyPheArgIleArgCysCysSerPhe 252
395 ..... 395
252 roGlyPheArgAsnArgHis**IleLeuArgCysPheHisCysArgTyr 268
394 .....CCCATGATG 384
269 **SerCysArgArgCysArgCysProArgCysPheGlyCysArgGlyCys 285
383 CGGT...AAGGCTGATCCGTCACAGGACTACGCTTCATCAGATC 337
285 sarGlyGlnGlyCys**SerHisArgArgPheArgCysArgGlyCysC 302
336 GGAATGTGAG.....CATGCTGTGCAAGGGGGAATGACTT 299
302 yAsnCysArgCysTrpArgCysArgGlyCysSerArgArgProGlyLeu 318
298 CGTCCCGCTGAT.....CGGAAGCGGACATGT.. 269
319 ProGlyArgAsp**ArgProValGlyHisArgGlyIleProthnCysCy 335
268 .....AGCCGATATTTCTTAAATGCGCGC 244
335 sPheArgCys**ArgSerProArgSerArgProAlaLeu**TrpPro 352
243 CTGCTGATGATGACAGTTGCCCACTGATGCTTGTATGTTCCCAT. 195
|||::: |||::: |||::: |||
352 roGly.....SerCys**ThrAsnProIleArgCysCysProser 365
194 .....CCGATATGACCGCTGCGCTCGGCAACTTCCCTCTCTCC 154
366 ***SerArgProIle...ProAlaArg.....ProArgLeuP 377
153 GAATGCTGTATTTCCGCTGCGGTGCAATGCTGACGTCGAGACCT 104
377 ogIyArgSerTyrArgTyrProProThrIlySerIlyArgSerGlnAsnC 394
103 GCCGATAAAGATCGTTTGCCAAATCT 75
394 ystRphIsArgSerSerGlySerArgThr 403
seq_name: /cgn2_6/ptodata/1/1aa/5b.COMB.pep:US-08-616-844-40
seq_documentation_block:
Sequence 40, Application US/08616844
Patent No. 5849578
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A.
TITLE OF INVENTION: COMPOSITION AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESS: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/616,844
 FILING DATE: 15-MAR-1996
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/599,654
 FILING DATE: 09-FEB-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/485,573
 FILING DATE: 07-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/386,844
 FILING DATE: 10-FEB-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: CORUZZI, LAURA A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7853-053
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 40:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1481 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-616-844-40

alignment_scores:
 Quality: 110.50 Length: 380
 Ratio: 0.617 Gaps: 17
 Percent Similarity: 47.105 Percent Identity: 20.263

alignment_block:
 US-09-303-518D-465 x US-08-616-844-40 ..

Align seg 1/1 to: US-08-616-844-40 from: 1 to: 1481

```

167 AACCTCCGACGCGCATATGCGATGGAACATACAAACCAT 216
    ||| |||: |||: |||: |||: |||: |||: |||: |||
674 AsnSerSerSerSerSerSerSerSerSerSerSerSer 690
    ||| |||: |||: |||: |||: |||: |||: |||: |||
217 CAGT.....TGGCAACCTGTTTCATCCAGCGCGC 248
    ||| |||: |||: |||: |||: |||: |||: |||: |||
690 eSerAsnSerSerSerSerSerSerSerSerSerSerSer 707
    ||| |||: |||: |||: |||: |||: |||: |||: |||
249 CATTAAAGAAATGCGGTACATTTGCCGCTTT..... 283
    ||| |||: |||: |||: |||: |||: |||: |||: |||
707 IuArgSerSerSerSerSerSerSerSerSerSerSerSer 723
    ||| |||: |||: |||: |||: |||: |||: |||: |||
284 .....CCGATCAGGCGCAGAGTCATCCATCC.....CCTTGCAC 318
    ||| |||: |||: |||: |||: |||: |||: |||: |||
724 GluSerProValLeuHisThrSerAsnLeuProSerTyrThrProThrI 740
    ||| |||: |||: |||: |||: |||: |||: |||: |||
319 AACCATGCTCACATTCGATTCTGATGAAGCGGTAAGTCCGCTGACCG 368
    ||| |||: |||: |||: |||: |||: |||: |||: |||
740 eAsnMetProAsnThrSerValValLeuAsp.....ThrA 752
    ||| |||: |||: |||: |||: |||: |||: |||: |||
369 ATTACGCTTTACCGCATTCATGGGAGAGATAGACACCATCCCGCG 418
    ||| |||: |||: |||: |||: |||: |||: |||: |||
752 sPAlaIuPheValSerAspSerSerSerSerSerSerSerSerSer 768
    ||| |||: |||: |||: |||: |||: |||: |||: |||
419 ACGGCTATGACGGCCACAGCGCGGCTATCCCGTCCCAAGGCGCG 468
    ||| |||: |||: |||: |||: |||: |||: |||: |||
769 SerSerSerSerGlyPro.....ProLeuPro..... 777
    ||| |||: |||: |||: |||: |||: |||: |||: |||
469 AGGATATATACAGCTACGACATAAAGCGCTGCCCAAAATATCC.... 514
    ||| |||: |||: |||: |||: |||: |||: |||: |||
778 .....GCTCAACCTGACCGACAAACCGACGACCGGACGACAC 550
    ||| |||: |||: |||: |||: |||: |||: |||: |||

```

```

784 eRhHisIleuPheSerSerIleLeuProSerThrArgAlaSerValHis 800
    ||| |||: |||: |||: |||: |||: |||: |||: |||
551 GCGTTGTCGACCGCTTCCACAATACCGGTAGTAGCTGACCAAGAGCTA 600
    ||| |||: |||: |||: |||: |||: |||: |||: |||
801 LeuLeuLysSerThrSer..... 806
    ||| |||: |||: |||: |||: |||: |||: |||: |||
601 GCGCAGCATTCACAGCGCCACCGCATACAGCCCGAGTGCAGACATC 650
    ||| |||: |||: |||: |||: |||: |||: |||: |||
807 .....AspAlaSerThrProTyrPseSerSerProSer..... 817
    ||| |||: |||: |||: |||: |||: |||: |||: |||
651 GGGCAATGCGCGCCGAGGCTTTCACGCGCATGCAGATTCGTCAAAACA 700
    ||| |||: |||: |||: |||: |||: |||: |||: |||
818 .....ProLeuProValSerLeuThrThr 825
    ||| |||: |||: |||: |||: |||: |||: |||: |||
701 TCATGCGCGCGCAGAGAAATTCGCGCGCAGCGATCCGCTGACAGGT 750
    ||| |||: |||: |||: |||: |||: |||: |||: |||
826 SerThrSerAlaProLeuSerValSerIleThrThrLeuPro..... 839
    ||| |||: |||: |||: |||: |||: |||: |||: |||
751 ATAAAGCAAGGCTCAACATTCGTGTATGACGCGCTGGGTGCTTTC 800
    ||| |||: |||: |||: |||: |||: |||: |||: |||
840 .....GlnSerSerSerThrProValLeuProArgAlaArgGlu.....T 853
    ||| |||: |||: |||: |||: |||: |||: |||: |||
801 CACGCAAAACAGATGCGCGCATCAGCATTCGCGCAGATATGCGCGC... 847
    ||| |||: |||: |||: |||: |||: |||: |||: |||
853 hrProValThrSerPheGlnThrSerThrMetThrSerPheMetThrMet 869
    ||| |||: |||: |||: |||: |||: |||: |||: |||
848 .....AACTCAAAAGACTATGCGCGCAGCA 870
    ||| |||: |||: |||: |||: |||: |||: |||: |||
870 LeuHisSerSerGlnThrAlaAspLeuLysSerIleThrThrProHisG 886
    ||| |||: |||: |||: |||: |||: |||: |||: |||
871 GCCATCCGCGATTTGGCAGTCACAAACCCCA.....ATGCGCGC 908
    ||| |||: |||: |||: |||: |||: |||: |||: |||
886 nGluLysValIleThrGlnSerLysSerProSerLeuValSerLeuProT 903
    ||| |||: |||: |||: |||: |||: |||: |||: |||
909 ACAAGGCATAGACGCGTCACCAATATCTTACGCGCATATCCCGCTCA 958
    ||| |||: |||: |||: |||: |||: |||: |||: |||
903 hrGlnSerThrLys.....AlaValThrThrAsnSerProLeuProPro 917
    ||| |||: |||: |||: |||: |||: |||: |||: |||
959 AAGGATTTGAGAGCTGTTGGGGAATACGCGCTGGCGGCGCATACGCA 1008
    ||| |||: |||: |||: |||: |||: |||: |||: |||
918 SerLeuThrGlnSerSerThrGlnGlnThrLeuProAlaThrSerThrAs 934
    ||| |||: |||: |||: |||: |||: |||: |||: |||
1009 CATCCTGTCAAGCGCGTCAGATGCGCGAGATCGCATTCGCGAAAGGAA 1058
    ||| |||: |||: |||: |||: |||: |||: |||: |||
934 nLeuAlaGlnMetSerProThrPhe..... 942
    ||| |||: |||: |||: |||: |||: |||: |||: |||
1059 ATCCGCGCTACGCACAATTTG.....CCGATGCGGCAATAGC 1096
    ||| |||: |||: |||: |||: |||: |||: |||: |||
943 .....ThrThrThrIleLeuLysThrSerGlnProLeuMetThrThr 956
    ||| |||: |||: |||: |||: |||: |||: |||: |||
1097 CCA.....AATACCCGTCCTTACCATTCGCAAAATATTCCT 1134
    ||| |||: |||: |||: |||: |||: |||: |||: |||
957 ProGlyThrLeuSerSerThrAlaSerLeuValThrGlyProIleAlaVal 973
    ||| |||: |||: |||: |||: |||: |||: |||: |||
1135 TCAAACTTGGAGCAGCGCTTACGCGCAAAAGAAACATCACT 1174
    ||| |||: |||: |||: |||: |||: |||: |||: |||
973 IGIlnThrThrAlaGlyGlnSerLeuThrThrHisPro 986
    ||| |||: |||: |||: |||: |||: |||: |||: |||

```

seq_name: /cgn2.6/ptodata/1/1aa/5B.COMB.pep:US-08-599-654-40

seq_documentation_block:
 ; Sequence 40, Application US/08599654
 ; Patent No. 5882925
 ; GENERAL INFORMATION:
 ; APPLICANT: PALB, DEAN A
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 ; TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
 ; NUMBER OF SEQUENCES: 54
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York


```

; seq.docmentation_block:
; Sequence 40. Application US/08944423A
; Patent No. 6020463
;
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,423A
; FILING DATE: 06-OCT-1997
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: JUN-07-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
;
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
;
; US-08-944-423A-40

```

```

284 .....CGATCCAGGGGACGAACTGCATTCGCC.....CCTCGAC 318
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
724 GluSerProValLeuHisThrSerAlaLeuProSerThrProHis1 740
319 AACATGCGCTCAACTTCGATTCGTGATGAAGACCGGTAGTCCCTGTGACG 368
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
740 eAsmMetProAlaSerHisThrSerValAlaLeuAsp.....Thr 752
369 ATTGACGCTTTACGCGATTCATGTGGACGGATACGAACACCATCCGCCG 418
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
752 sPaLaGluPheValSerAspSerSerSerSerSerSerSerSerSer 768
419 ACGGCTATGACGGGGCGACAGGCGCGCGGTATCCGCGTCCCAAAAGCGCG 468
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
769 SerSerSerSerGlyPro.....ProLeuPro..... 777
469 AGGGATATATACGTACGTACGACATATAAAGCGCTTGCCCAAAATATAC... 514
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
778 .....LeuProSerValSerLns 784
515 .....GCTCAACCTGACCCGACACACGACGACGACGACGAC 550
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
784 eRHisHisLeuPheSerSerLeuLeuProSerThrArgAlaSerValHis 800
551 GCGTTGTGACCGCTTCCACAAATACCGGTATATGTCGACGACGAGTA 600
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
801 LeuLeuLysSerThrSer..... 806
601 GCGCAGCGATTCCAAAAGCGCGCACCGGATACAGCCCGACGTGACAGATC 650
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
807 .....AspAlaSerThrProTyrPserSerSerProSer..... 817
651 GGGCAATGCGCGCGGACGCTTTCACACGGCAGTCGCAAGATGCTCAAAACA 700
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
818 .....ProLeuProValSerLeuThrThr 825
701 TCATGCGCGCGCGGACAGACAAATGTGCGCGGACGAGCGATGCCGTGCAGGT 750
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
826 SerThrSerAlaProLeuSerValSerLnsThrHisLeuPro..... 839
751 ATAACGAGAAAGGCTCAAACTGTGTTATGACAGCGCTTGCGTCTGTTTC 800
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
840 .....GlnSerSerSerThrProValLeuProArgAlaArgLys.....T 853
801 CACCGAAAAACAAGATGGCGCGCATCAACAGATTTGGCAGATATGGCGC... 847
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
853 hProValThrSerPheGlnThrSerThrMetThrSerPheMetThrMet 869
848 .....AATCTCAAAAGACTATGCCGCGACGA 870
870 LeuHisSerSerGlnThrAlaAspLeuLysSerGlnSerLnsThrProHisG1 886
871 GGCATCCGGGATTTGGGACAGTCCAAACCCCA.....ATGCCGC 908
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
886 nGluLysValIleThrGlnSerLysSerProSerLeuValSerLeuProT 903
909 ACAAGCGCATAGAACCGCTGACGCAATATGTTTACGCGCATCTCCCGCTCA 958
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
903 hGlnSerThrLys.....AlaValThrThrAsnSerProLeuProPro 917
959 AAGGATTTGGACCTGTTGGGGAATAATACGCTTGGCGGCATCAAGCA 1008
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
918 SerLeuThrGlnSerSerThrGlnGlnThrLeuProAlaThrSerThrAs 934
1009 CATCCCTGTCAACCGGTGCGAGATGGCGAGATCGCATTTGCCGAAGAAGGAA 1058
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
934 nLeuAlaGlnMetSerProThrPhe..... 942
1059 ATCCGCGCGTCCAGCAGCAATTTTG.....CCGATCCGGCATACG 1096
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
943 .....ThrThrThrIleLeuLysThrSerGlnProLeuMetThrThr 956
1097 CCA.....ATAACCGCGTCCCTTACCATTTCCGGAATAATATCCGT 1134

```


TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
FEATURE:
NAME/KEY: UNSURE
LOCATION: (4)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (6)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (20)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (29)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (54)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (64)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (69)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (89)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (99)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (114)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (119)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (129)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (139)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (169)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (182)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (185)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (219)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (259)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (269)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (291)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (323)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (339)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (349)
NAME/KEY: UNSURE

LOCATION: (356)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (366)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (410)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (421)
US-09-997-897-2

alignment_scores:
Quality: 116.50 Length: 343
Ratio: 0.932 Gaps: 19
Percent Similarity: 36.443 Percent Identity: 25.948

alignment_block:

US-09-303-518d-465/rev x US-08-997-897-2 ..

Align seg 1/1 to: US-08-997-897-2 from: 1 to: 430

888 TGCCCCAATCGCGGATGCTGCTGCGGCAATGCTTTGAGTTGCCCATAT 839
|||||
121 CysPro.....CysCysSer..... 125
838 CGCCCAATCGTTGATGCGCGCCATCTTTGCGTGAAGACAGAC.C 790
|||||
126 .CysGlnHis**GlyCysArgTyrCysArgTyrProGlySerArgTyrP 142
789 CAGCCGTCATACGACGAAATGTTGAGCTTCGCTTATACCTGCACGG 740
|||||
142 roser.....SerArgCysProSerLeuArgCys..... 151
739 CATGCCCTGCCGCCACATTTCTCTGCGCGCCGATGATGTTTACG 690
|||||
152ArgArgPheArgCysProArg**ArgCys..... 161
689 ATATCTGCGAGTCCGTTGAAGCTTCGAGCGCATTCGCCATCTGTCAG 640
162GlnArgTyrTrpCysProAsn**ThrG 171
639 CTCGGGCTGTATCGGGTGGCGGCTTTGAATCGTCCTTACTCTTGGC 590
:::|||||
171 LysArgCysCys.....ArgCysProSerSer 179
589 TCA.....GCATCTACCGGTATTTGGAACGCTGACACAGCCGTTGT 546
|||
180 SerArg**GlnTyr**SerProAlaGlyCysArgArgThrAlaArgCys 196
545 CCGGTGCTGCGGTGTCGTCAGGTGAGGGGATATTTGGGCAAGCC 496
|||||
196 SarArgCysCysArgCys.....TyrArgT 206
495 TTTATGTCGTAGCTGTATATATCCCTCGCGCCTTTGGAG..... 455
|||
206 hrArgCysCysCys.....ArgCysTrpGlnSerLeuGly 218
454CGGATAGCCGCCCTGTGGC 432
219 **SerArgProArgSerArgSerArgArgCysSerArgArgArgPheG1 235
431 CCGTCATAGCGGTGCGGGATGTTGCTATCCGT..... 395
235 nAsnArgCysArgSerArgGlyPheArgIleArgCysCysSerPheP 252
395 395
252 roGlyPheArgAsnArgHis**IleLeuArgCysPheHisCysArgTyr 268
394CCCAATGAGAG 384
:::|||||

```

269 ***SerCysArgArgCysArgArgCysPheIlyCysArgIlyCys 285
383 CGGT...AAGGCTGAATCCGTACACGGGACCTACGGCTTCATCAGAAATC 337
285 sarGcysglIyCys***serIsarGArgPheArgCysArgIlyCysC 302
336 GGAATGTAGG.....CATGGTTGTCAAGGGGGAATGACTT 299
302 ysAsnCysArgCysTrpArgCysArgIlyCysSerArgArgProglIyLeu 318
298 CGAGCCCGCTAT.....CGCAAAAGCGCAAAATCT.. 269
319 ProglIyArgAsp***ArgProValglIyHisArgIlyLeuProthrCysCys 335
268 .....AGCGATATTTCTTATGGCGCG 244
335 sPheArgCys***ArgSerProArgSerArgProAlaLeu***TrpPro 352
243 CTGCTGGATGACAGATTGCCCACTGATGCTTTGATGTTTCCCAAT. 195
352 roglIy.....SerCys***ThrAsnProIleArgCysCysProser 365
194 .....CCGATATGACCGCTGCGCTCGCAAGTTCCCGCTGCTGCC 154
366 ***SerArgProIle...ProAlaArg.....ProArgLeuPr 377
153 GAATAGGTGATTTCCCGTGGGTTGCAAAATGCTGACGGTGCAGAACT 104
377 OGlyArgSerTrpArgTrpProProthrIlySerGlyArgSerIlnsnc 354
103 GCCGATTAAGAAATCGTTGCCAAATCT 75
394 ystrpHisArgSerSerGlySerArgThr 403

```

seq_name: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:US-09-035-648-24

seq_documentation_block:

; Sequence 24, Application US/09035648

; Patent No. 6100031

; GENERAL INFORMATION:

; APPLICANT: Shyjan, Andrew W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; TITLE OF INVENTION: DIAGNOSIS, PREVENTION, AND TREATMENT OF NEOPLASTIC CELL

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson, P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: PastSEO for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/035,648

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/818,829

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Melike John, Ph.D., Anita L.

; REGISTRATION NUMBER: 35,283

; REFERENCE/DOCKET NUMBER: 07334/003001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-542-5070

; TELEFAX: 617-542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 24:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 739 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-09-035-648-24

```

alignment_scores:

```

Quality: 111.50 Length: 528
Ratio: 0.489 Gaps: 28
Percent similarity: 43.182 Percent identity: 21.591

```

alignment_block:

US-09-303-518d-465 x US-09-035-648-24

Align seg 1/1 to: US-09-035-648-24 from: 1 to: 739

```

227 ACCGTTCATCCAGCAGCGGCCCATTAAGAAATATGCGTACATTTGTC 276
111 : : : : : : : : : : : : : : : : : : : : : : : : : :
178 rhrThrThrSerTyrGlnArgIlyGlnIlyGlnIlySerValArgThrClse 194
277 GCCTTTCCGATCAGCGCAGCAAGTCC..ATTCCTTCGACACCA 323
194 rSerVal.....SerIlySerTyrIleGlnValSerHisGlyL 207
324 TGCTTCACATTCGG.....ATT 340
207 euGlyHisValProGlnGlnAsnGlyPheSerGlyAlaSerCluThrAla 223
341 CTGATGAAGCCGCTAGTCCGTTGACAGATTCAGCC.....TTTACCGC 384
224 AsnIleGlnArgMetValProMetAsnArgThrAlaIlyAsnProPheProgl 240
385 ATTCATTTGGGACGATACGACACCATCCCGCGCAGGCTATGACGGCC 434
240 ySerValAsnGlnProAlaThrProPheSerProThrArgAsnMetThrS 257
435 ACAAGGCGGCGGCTATCCCGCTCCCAAGCGCGCAGGGATATATACACT 484
257 eTrpIleAlaAspPheProAlaPro..... 265
485 ACGACATAAAGCGTTGCCCAAAATATCCGCTACCTGACCGCAGCAAC 534
266 .....ProProIlySerAlaValThrProPro..... 274
535 CGCAGCACCAGCAACAGCGCTTGTGACGCTTTCCACATACCGGTATAT 584
275 .....ProAspAlaPheSerArgGlyValSerSer...ProIle.... 286
585 GCTGACGCAAGGATGAGCGGAGGATTCAAAGCGCGCACCGCATACAGCC 634
287 .....AlaGlyProAlaIlePro 292
635 CCGAGCTGACAGATCGGCAATGCCCGCAAGCTTTCACAGCGCACTGCA 684
293 ProProThrProGlnProAla.....ProTrpSerGlnProAlaPheIly 307
685 GATATGCTCAAAAACATCATTCGCGCGCGCAGAGAAATTTGCGCGCAG 734
307 rAspSerSerGluArgIleAlaSerArgAspGlnArgIleSerValProA 324
735 CGATGCCGTGACAGGATATAAGGAAGGCTCAAAACATTTGCTGTATGACG 784
324 lAllys..... 325
785 GCTTGGGTCTGCTTTCACCGAAACAGATGCGCGCATCAGCATTTG 834
326 .....ArgThrGlyIleLeucl 331
835 GCAGATATGCGCGCACTCAAAAGCTATGCCG.....CAGCAGC 872
: : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

331 ngluAlaLysArgSerThrThrlsPrometPheThPhelysGluP 348
873 CATCCGGATTGGGACAGTCCAAACCCCAATGCCGACACAGGATAGAG 922
348 rplysVal.....SerPro.....Asn 353
923 CCGTCACCAATATCTTACGGCAGTCATCCCGCTCAAG...GGATTGA 969
354 ProLeuLeuSerLeuLeuGlnAsnSerGluGlyLysArgGlyThrG1 370
970 GCTGTGGGGAATACGGCTTGGGGGATACGGCAGCATCTCTGCA 1019
370 yAlaGlyLysPser.....GlyProGluGlnAspTyrLeuSerL 384
1020 GCGGTCCGACAG.....TGGCGAGATCGCATTCGCCGAAGAGA 1057
384 euglyAlaGluAlaCysAsnPhetClnSerSerSerAlaLysGlnLys 400
1058 AATCCGCGCTACGACGACATTTTGGCGATGCGGACATACGCAATACCCG 1107
401 ThrProProValAlaProLysProAlaValLysSerSerSerG1 417
1108 TCCCTTACCATTCGCAATATCCGTTCAACTTGGAGACGGTTACG 1157
417 nProValThr...ProValSerProVal.....TTPSerProGlyValA 431
1158 CAAGAAACATGACCTCTCTCA.....ACCGTGGCGCGCTCAACGGA 1200
431 La...ProThrGlnProProAlaPheProThrSerAsnProSerLysGly 446
1201 .....AAGATGTGAACCTGGCAACAAACGCCACCGCAAGACCA 1241
447 ThrValValSerSerLeuLysIleAlaGlnProSerTyrProProAla 463
1242 AGTG.....CCGTTGACCGGAAGGCTTC 1267
463 gProAlaSerThrLeuAsnValAlaGlyProPhe.....LysGlyProG 478
1268 CGAATTTTGAAGAAAGAGTAAATACGATACGAGATTAATACCGCTGA 1317
478 lmlAlaValAlaSerGlnAsnTyrThrProLysProThrValSerThr 494
1318 CCACAGTGAATCTTATAGATGACCGCTCTTAATCTTAAGGTTCTGT 1367
495 ProThrValAsnAlaValGlnProGlyAlaValGlyProSerAsnGlu 511
1368 CGGATCG..... 1374
511 uProGlyMetSerGlyArgGlyAlaGlnLeuPheAlaLysArgGlnSerA 528
1375 .....GCTCATTTCTTGGCTATTAAGTCCAGAAATTCAA 1407
528 rGmetGluLysTyrValValAlaPserAspThrValGlnlmlAsnAla 544
1408 TACGCCAAATTTACCAAG.....CAAG 1430
545 ArgAlaGlnSerProThrProSerLeuProAlaSerTyrLysTyrSer 561
1431 TAGAATCAGATATATCCACCTTAAATTAAC.....TCGCCTT 1468
561 rAsnValAlaGlnAlaProProValAlaTyrAsnProIleHisSerPro 578
1469 CAGCACCG.....CTACCAAAAGACCTAATTAAGATATTGGAT 1509
578 eTyrProLeuAlaAlaLeuLysSerGlnProSerAlaAlaGlnProSer 594
1510 AATTTGGTAATGATGATGAAGTCCATCAAGACT..... 1548
595 LysMetGlyLysLysLysGlyLysProLeuAsnAlaLeuAspValMe 611
1549 ..AAGGTCAAGATTTGAATGGGATGTCAATTG 1581
611 tlySHISGlnProTyrGlnLeuAsnAlaSerLeu 622

```

```

seq_name: /cgn2_6/plodata/1/laa/6B_COMB pep:US-09-001-951-24
seq_documentation_block:
; Sequence 24, Application US/09001951
; Patent No. 6268470
; GENERAL INFORMATION:
; APPLICANT: Shyan, Andrew W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: DIAGNOSIS, PREVENTION, AND TREATMENT OF NEOPLASTIC CELL
; TITLE OF INVENTION: GROWTH AND PROLIFERATION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/001,951
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/818,829
; FILING DATE: 14-MAR-1997
; APPLICATION NUMBER: 60/013,438
; FILING DATE: 15-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Melkielejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 739 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-09-001-951-24

```

```

alignment_scores:
Quality: 111.50 Length: 528
Ratio: 0.489 Gaps: 28
Percent Similarity: 43.182 Percent Identity: 21.591

```

alignment_block:

US-09-303-518d-465 x US-09-001-951-24 ..

Align seg 1/1 to: US-09-001-951-24 from: 1 to: 739

```

227 ACCTGTCATCCAGCAGCGCATTAAGGAATATCGCTACATTGTC 276
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
178 ThrThrThrSerTyrGlnArgLysGluGlnGluSerValArgThrGln 194
277 GCGTTTTCGATCAAGGCGACGAAGTCC...ATTCCCCCTTCGACAACA 323
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
194 rSerVal.....SerLysSerTyrIleGluValSerHisGlyL 207
324 TGCCTCACATTCG.....ATT 340
207 euglyHnValProGlnGlnAsnGlyPheSerGlyAlaSerGluThrAla 223

```

```

341 CTGATTAAGCCGGTAGTCCCGTTGACGATTCAGCC.....TTTACCGC 384
      : : : : : : : : : : : : : : : : : : : : : : : : : :
224 AsnIleGlnArgMetValProMetAsnArgThrAlaLysProPheProcl 240
      : : : : : : : : : : : : : : : : : : : : : : : : : :
385 ATCCATGGGAGCGATACGACACCATCCCGCGGCGGTATGACGGGCC 434
      : : : : : : : : : : : : : : : : : : : : : : : : : :
240 ySerValAsnGlnProAlaThrProPheSerProThrArgAsnMetThrS 257
      : : : : : : : : : : : : : : : : : : : : : : : : : :
435 ACAGGGCGGCGGTATCCCGTCCCAAGAGCGGAGGATATATACAGT 484
      : : : : : : : : : : : : : : : : : : : : : : : : : :
257 erProIleAlaAspPheProAlaPro..... 265
      : : : : : : : : : : : : : : : : : : : : : : : : : :
485 ACGACATTAAGAGCGTTGCCCAAAATATCCGCTCAACCTGACGCAAC 534
      : : : : : : : : : : : : : : : : : : : : : : : : : :
266 .....ProProTyrSerAlaValThrProPro..... 274
      : : : : : : : : : : : : : : : : : : : : : : : : : :
535 CGGAGACCGGAGCAAGCGCTTGTCGACCGCTTTCACAAATACCGGTAT 584
      : : : : : : : : : : : : : : : : : : : : : : : : : :
275 .....ProAspAlaPheSerArgLysAlaSerSer.....ProIle.... 286
      : : : : : : : : : : : : : : : : : : : : : : : : : :
585 GCTGACGCAAGAGTAGGCGAGGATTCAAAGCGCGCCCGATACAGCC 634
      : : : : : : : : : : : : : : : : : : : : : : : : : :
287 .....AlaGlyProAlaGlnPro 292
      : : : : : : : : : : : : : : : : : : : : : : : : : :
635 CCGAGCTGACAGATCGGCAATCGCGCGAGCTTTCACGCGCACTGCA 684
      : : : : : : : : : : : : : : : : : : : : : : : : : :
293 ProProTyrProGlnProAla.....ProTyrSerGlnProAlaPheTyr 307
      : : : : : : : : : : : : : : : : : : : : : : : : : :
685 GATATGCTGAAAAACATCATCGCGCGGAGGAGGAAATGTGCGCGCAG 734
      : : : : : : : : : : : : : : : : : : : : : : : : : :
307 rAspSerSerGlnArgLysAlaSerArgAspGlnArgLysSerValProA 324
      : : : : : : : : : : : : : : : : : : : : : : : : : :
735 CGATGCCGTGACAGGATATAGCAAGGCTCAAAATGCTGTATGACAG 784
      : : : : : : : : : : : : : : : : : : : : : : : : : :
324 LaLys..... 325
      : : : : : : : : : : : : : : : : : : : : : : : : : :
785 GCTGGGCTGCTTCCACCGAAACAGATGCGCGCATCAAGATTGG 834
      : : : : : : : : : : : : : : : : : : : : : : : : : :
326 .....ArgThrGlyIleLeuGln 331
      : : : : : : : : : : : : : : : : : : : : : : : : : :
835 GCAGATATGGCGCACTCAAGACTATGCG.....CAGCAC 872
      : : : : : : : : : : : : : : : : : : : : : : : : : :
331 ngAlaLysArgArgSerThrThrLysProMetPheThrPheLysGlnP 348
      : : : : : : : : : : : : : : : : : : : : : : : : : :
873 CATCCCGCATTTGGGAGTCCCAAAACCCCATGCGCAGCAAGCATAGAG 922
      : : : : : : : : : : : : : : : : : : : : : : : : : :
348 roLysVal.....SerPro.....Asn 353
      : : : : : : : : : : : : : : : : : : : : : : : : : :
923 CCGTCAGCAATATCTTTACGCGATCATCCCGCTCAAG...GGATTGGA 969
      : : : : : : : : : : : : : : : : : : : : : : : : : :
354 ProGlnLeuLeuSerLeuLeuGlnAsnSerGlnGlyLysArgGlyThrG 370
      : : : : : : : : : : : : : : : : : : : : : : : : : :
970 GCTGTTGGGGAATATAGGCTTGGGCGCATCAGCGCATCTCTCAA 1019
      : : : : : : : : : : : : : : : : : : : : : : : : : :
370 yAlaGlyLysAspSer.....GlyProGlnGlnAspTyrLeuSerL 384
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1020 GCGGTCGCA.....TGGGCGAGATCGCATGCGCAAGGA 1057
      : : : : : : : : : : : : : : : : : : : : : : : : : :
384 euGlyAlaGlnLacysAsnPheMetGlnSerSerSerAlaLysGlnLys 400
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1058 AATCCCGCTCAGCGACAAATTTTCCGATGCGGATACGCCAAATACCG 1107
      : : : : : : : : : : : : : : : : : : : : : : : : : :
401 ThrProProProValAlaProLysProAlaValLysSerSerSerSerG 417
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1108 TCCCTTACCATTCGCGAAATATCGGTCAAACTTGAGCAGCGTTAGG 1157
      : : : : : : : : : : : : : : : : : : : : : : : : : :
417 nProValThr...ProValSerProVal.....TyrSerProGlyValA 431
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1158 CAAAGAAACATCACTCTCTCA.....ACGTCGCGCGCTCAACAGCA 1200
      : : : : : : : : : : : : : : : : : : : : : : : : : :
431 La...ProThrGlnProProAlaPheProThrSerAsnProSerLysGly 446
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1201 .....AAGCATGTGAACCTGGCAAAACAGCCCGCAAGACCA 1241

```

```

447 ThrValValSerSerIleLysIleAlaGlnProSerTyrProProAlaAr 463
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1242 AGTG.....CCGTTGACGAGTAAGGTTTC 1267
      : : : : : : : : : : : : : : : : : : : : : : : : : :
463 gProAlaSerThrLeuAsnValAlaGlyProPhe.....LysGlyProG 478
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1268 CGAATTTTGAAGAACGTAAATACGATACGAAATTAATACCGCTGTA 1317
      : : : : : : : : : : : : : : : : : : : : : : : : : :
478 InAlaAlaValAlaSerGlnAsnTyrThrProLysProThrValSerThr 494
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1318 CCACAGTAGTAATCCTATAGATGAACCCGCTTTAATCCTTAAGGTTCTG 1367
      : : : : : : : : : : : : : : : : : : : : : : : : : :
495 ProThrValAsnAlaValGlnProGlyAlaValGlyProSerAsnGlnLe 511
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1368 CGGATCG..... 1374
      : : : : : : : : : : : : : : : : : : : : : : : : : :
511 uProGlyMetSerGlyArgGlyAlaGlnLeuPheAlaLysArgGlnSerA 528
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1375 .....GCTCATCTTGCTGTATATACGCGCAGATTCAA 1407
      : : : : : : : : : : : : : : : : : : : : : : : : : :
528 rGMeGlnLysTyrValValAlaAspSerAspThrValGlnAlaHisAla 544
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1408 TACGCAAAATTAACAAG.....CAAG 1430
      : : : : : : : : : : : : : : : : : : : : : : : : : :
545 ArgAlaGlnSerProThrProSerLeuProAlaSerThrTyrLysSerSe 561
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1431 TAGAATCGATATATCCACCTAAAAATTAC.....TCTCCTT 1468
      : : : : : : : : : : : : : : : : : : : : : : : : : :
561 rAsnValAlaArgAlaProProProValAlaTyrAsnProIleHisSerSe 578
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1469 CAGCACG.....CTACCAAGAGACCTAATATGATATTTGAT 1509
      : : : : : : : : : : : : : : : : : : : : : : : : : :
578 erTyrProLeuAlaAlaLeuLysSerGlnProSerAlaAlaGlnProSer 594
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1510 AAATTGTAATGAATGACTAAGGTCATCAAGAACT..... 1548
      : : : : : : : : : : : : : : : : : : : : : : : : : :
595 LysMetGlyLysLysLysGlyLysLysProLeuAsnAlaLeuAspValMe 611
      : : : : : : : : : : : : : : : : : : : : : : : : : :
1549 .AAAGTCAGAAATTTGAATGGAGTTCATTTG 1581
      : : : : : : : : : : : : : : : : : : : : : : : : : :
611 tLysHisGlnProTyrGlnLeuAsnAlaSerLeu 622

```

seq_name: /cgn2_6/prodata/1/laa/5B_COMB.pep:US-09-367-206-20

seq_documentation_block:

```

; Sequence 20, Application US/09367206
; Patent No. 6326482
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: NSP Molecules
; FILE REFERENCE: P1223RIE
; CURRENT APPLICATION NUMBER: US/09/367,206
; CURRENT FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/08847
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 60/082,767
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 60/113,296
; PRIOR FILING DATE: 1996-12-22
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 20
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-367-206-20

```

alignment_scores:

Quality:	Length:
Ratio: 109.50	240
Percent Similarity: 0.928	Gaps: 14
Percent Identity: 49.167	Percent Identity: 25.833

alignment_block:

US-09-303-518D-465 x US-09-367-206-20 ..

Align seg 1/1 to: US-09-367-206-20 from: 1 to: 688

```

82 GCAACACATCTTTT.....ATCCGCGAGTTCCTGCACCGTCATGACATT 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
102 AlalgluSerThrThrHisIleGlnTyrLeuPheGlnGlnGlnSerPhe 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
126 CGAACCCGACGGG.....AATACGACCTATTGCGAGCAGCGGGG 166
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
118 eAspHisValProAlaLeuValArgTyrHisVal..GlySerArgLysA 134
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 AACTTGGCGAGCGCGGTCATATC.....TTCGCGCTTTTC 284
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
134 laValSerGlnGlnSerGlyAlaIleIleTyrCysProValAsnArgThr 150
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
193 .....GGATTGGGAACAT...AC 208
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 PheProLeuArgTyrLeuGlnAlaSerTyrGlyLeuGlyGlnGlySer 167
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
209 AAAGCATCATGTTGGCACTGTTTCATCCAGACGCGGCA..... 250
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 rLysProAla.....SerProValSerProSerGly..ProLysGly 181
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
251 .....TTAAGGAATATATGCTACA.....TTCGCGCTTTTC 284
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
181 rHisMetLysArgArgSerValThrMetThrAspGlyLeuThrAlaAspL 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
285 CGATACAGGCGACGACAGTCCATTCCTCCCTTCGACACCATGCT.... 328
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
198 ySValThrArgSerAspGlyCysProThrSerThrSerLeuProArgPro 214
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
329 .....CACATTCGATTC 342
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
215 ArgAspSerIleArgSerCysAlaLeuSerMetAspGlnIleProAsp 221
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
343 GATGACGCGGTAGTCCCTGAGCGATTCAGCTTATCCGATTCATG 392
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
231 uHisSerProMetSerProIleSerGlnSerProSerSerProAlaTyrS 248
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
393 GGAGCGATAGACACCC..ATCCCGCGACGCGTATGACGGCGCACAG 439
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
248 erThrValThrArgValHisAlaAlaProAlaIleProSerAlaThrAla 264
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
440 GCGGCGCTATCCGCTCCCAAGCGCGAGGATATATACAGTACAGAC 489
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
265 LeuProAlaSerProVal..AlaArgArgSerSerGlnProGlnLeu.... 279
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
490 ATAAAGCGCTTGCCCAAAATATCCGCTCA..... 521
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
280 .....CysProGlySerAlaProLysThrHisGlyLysSerAsp 292
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
522 .....CCTGACGACACCGCGACGCGACCAACAGCGCTTGTCGACGCT 565
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
293 LysGlyProHisThrSerProSerHisThrLeuGlyLysAlaSerProSe 309
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
566 TCCACATATACCGGTATGCTGACGCAAGAGTAGCGACGATTCAAA 615
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
309 rProSerLeuSerSerTyrSerAspProAspSerGlyHisTyrCysGlnL 326
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
616 CGGCGCACCGCATACAG 632
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
326 euGlnProProValArg 331
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq_name: /gcn2_6/ptodata/1/1aa/6A.COMB.pep:US-08-713-118-2

seq_documentation_block:

```

; Sequence 2, Application US/08713118
; Patent No. 6040436
; GENERAL INFORMATION:
; APPLICANT: Franco, Rodrigo
; APPLICANT: Sun Chen, Ai Ku

```

```

; APPLICANT: Sney, David J.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HUMAN NEURONAL
; TITLE OF INVENTION: CALCIUM CHANNEL SUBUNITS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,118
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mata, Elizabeth W.
; REGISTRATION NUMBER: 38,236
; REFERENCE/DOCKET NUMBER: ACC96-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-9540
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-713-118-2

```

alignment_scores:

```

Quality: 108.50 Length: 557
Ratio: 0.424 Gaps: 36
Percent Similarity: 45.961 Percent Identity: 21.903

```

alignment_block:

US-09-303-518D-465 x US-08-713-118-2 ..

Align seg 1/1 to: US-08-713-118-2 from: 1 to: 2337

```

108 TCTGCACGCTCGCATTCGAAACCGGAGGAATACACCTATTGCGCA 157
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1898 ThrGlnProAlaValLeuArgGlyAlaArgValPhe.....LeuArgG 1912
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
158 GCGAGGGGGAACCTGCGAGCGCAGCGCTCATATCGATTGGGAACATA 207
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1912 nLysSerSerThrSerLeuSer..AsnGly.....GlyAlaIle 1924
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
208 CAAGCCATCACTGTTGGCAACCTGTTCAATCCAGCAGCGCGCATTAAGG 257
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1925 GlnAsn.....GlnGlnSerGlyIleLysG 1933
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
258 AAATTCGGGCTAC...ATTGTCGCTTTTCGATCAGCAGCAGCAGAC 304
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1933 uSerValSerThrGlyThrGlnArgThrGlnAspAlaProHisGlnAla 1950
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 ATTCCCTCTTGAC...AACCATGCTTCATTCGATTCTGATGAAACCC 351
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1950 rGProProLeuGlnArgGlyHisSerThrGlnIleProValGlyArgSer 1966
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1967 GlyAlaLeuAlaValAspValGlnMetGlnSerIleThrArgArgGlyPr 1983
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
402 CGAACACCATCCCGCGGCGATGATGAGCGGCGCACAGCGCGGCTATC 451
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1983 oAspGlyGlnProGlnProGly..LeuGlnSerGlnGlyArgAlaAla... 1998

```



```

452 CCGCTCCAAAGCGCGAGGATATATACAGTACACATATAAGCGTT 501
1999 .....:Serine 2000
502 GCCCAAAATATCCGCTCAACCTGACCGCAAC..GCAGACCG...G 545
1999 .....:Serine 2000
2000 tProatgLeuAlaAlaIaGlutHrGlnProValThraSPAlaSerPrometL 2017
546 ACAAGGGCTTGTGACCGTTTCCACAATACG.....GTAGTATGC 586
2017 ysaTgSerIleSerThrLeuAlaGlnaTgProaTgLyThrHisLeucys 2033
587 TGACGCAAGAGTAGGCGAGCATCAACGCCACCGATACAGCCG 636
2034 .....:SerThrThr 2037
637 GAGCTGACAGATCGGGCAATGCCCGAGCTTTCACGCGCATGCGACA 686
2037 O.....AspargProProProSerGlnAlaSerSer.His..... 2048
687 TATGTCMAAAACATCATCGGCGCGGCGAGAGAAATGTGCGCGCAGCGG 736
2049 .....HisHisHisHisArg.....CysHisArgTgTgAr 2058
737 ATGCGCTGAGGGTATATACGCAAGGCTCAACATTCGTGTATGACGCGC 786
2058 gaSPaTgLyS.....GlnaTgSerLeuLulys..... 2067
787 TTGGGTCTGCTTTCACCGAAACAAAGATGGCGCGCATACAGATT... 833
2068 ..GlyProSerLeuSerAlaaspMetaspGlyAlaProSerSerAlaVal 2083
834 .....GGCAGATATGGCGCGCAACTCAAAAGACT 859
2084 GlyProGlyLeuProProGlyGlnGlyProHnGly..... 2095
860 ATGCGCGCAGACCCATCCGATTTGGGCGAGTCCAAACCCCAATGCCGCA 909
2096 .CysArgaTgGlnaTgGlnaTg.....ArgGlnGlnaTgS 2107
910 CAAGCATATGAAAGCCGCTCAACATATCTTTACGGCAGTCATCCCGTCAA 959
2107 eArTgSerGlnGlnaTgArgTgProSerSerSerSerGlnuTgSgln 2123
960 AGGATATGAGCTGTCCGGGAAATACGGCTTGGCGCGC..... 999
2124 ArgPheTyrSerCysasp.....Arg.PheGlyGlyArgTgGluProP 2137
1000 .....ATCAGCGCACATCTGTCAAGGGTCCGCGAGT 1032
2137 rOlYsProLYsProSerLeuSerSerHisProThSerProThralaGly 2153
1033 GCGGAGATCGCA...TTGCCGAAGGAATCCGCGTCCGACAGCAATTT 1079
2154 GlnGlnaTgProGlnaTgProGlnaTgSerGlySerValasnGly..... 2168
1080 TGCGCATGGGCGCATAGCCCAATACCGCGCTTACCTCCCGAATA 1129
2169 .....SerProLeuLeuSerThrSer. 2175
1130 TCCGTTCAACTTGGAGCAGCGTTACGCAAAAGAAACATCACCTCCCTCA 1179
2176 ..GlyAlaSerThrProGlyArgGlyGlyArgTgGlnaTgProGlnaTg 2191
1180 ACCGTGCCCGCTCAACGCAAGAAATGTGAACCTGCAACCAACGCCCA 1229
2192 ProLeuThrProArgProSerIleThrTgLyTlnAlaasn..... 2205
1230 CCCGAAGACCAAGTCCGTTGACGGT.....AAAGGTTTCCGAATT 1273
2206 ....SerSerProIleHisPheAlaGlyAlaGlnThrSerLeuProAlaP 2221

```

```

1274 TTGAAAAAGAC...GTAAATATACGATACGAAATTAATACCGCTGACCA 1320
2221 heserProGlyArgLeuSerArgGlyLeuSerGlnHisAsnAlaLeuLeu 2237
1321 CAAGTGAATCCTATAGATGAACCGCTCTTAATCTTAAGAGTCTGCGG 1370
2238 GlnaTgAspProLeuSerGlnProLeu...AlaProGlySerArgIleG 2253
1371 ATCGGCTCAT.....TCTTGCTCTATACGCCGCAATTCATACCCAA 1414
2253 ySerAspProTgLeuGlyGlnaTgLeuAspSerGlnaTgAlaHisA 2270
1415 AATTACCA..... 1422
2270 lAleuProGlnaTgPThrLeuThrPheGlnGlnaTgAlaValaTlnAsnSer 2286
1423 ..AGCAAGTAGAATCAGATATATCCACCTAAATAATCTCTCTTC 1469
2287 GlyArgSerSerArgTgThrSerTgValSerSerLeuThr...SerGln 2302
1470 AGCACCGCTACCAAAAGACCTATATATGATTTGATTAATTTGCTA 1519
2302 HisProLeuArgTgAlaPro...AsnGlyTgHisCysThrLeuGly. 2317
1520 ATGAATGGACTAAAGTCCATCAAGAACTAAAGSTCAAGAAATTTGAATGG 1569
2317 ..... 2317
1570 GATGTTCAATTGTCTAAACAGAGAGACCACTTGAGTGGCTAGT 1619
2318 .....LeuSerSerGlyGlyArgAlaArgHisSerTgHisHisPr 2331
1620 GGATGCTAAGCAT 1632
2331 aspSPeLlnaspHis 2335

```

seq_name: /cgn2_6/prodata/1/laa/5B_COMB.pep:US-09-452-007-2

seq_documentation block:

```

; Sequence 2, Application US/09452007
; Patent No. 6140485
; GENERAL INFORMATION:
; APPLICANT: Franco, Rodrigo
; APPLICANT: Sun Chen, Ai Ru
; APPLICANT: Suey, David J.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HUMAN NEURONAL
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/452,007
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,118
; FILING DATE: 16-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mata, Elizabeth W.
; REGISTRATION NUMBER: 38,236
; REFERENCE/DOCKET NUMBER: ACC96-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240

```

TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2337 amino acids
; TYPE: amino acid
; TOPOLOGY: 1linear
; MOLECULE TYPE: protein
; US-09-452-007-2

alignment_scores:
Quality: 108.50 Length: 557
Ratio: 0.424 Gaps: 36
Percent Similarity: 45.961 Percent Identity: 21.903

alignment block:
US-09-303-518D-465 x US-09-452-007-2 ..

Align seg 1/1 to: US-09-452-007-2 from: 1 to: 2337

108 TCTGACCGTCAGCATTTGCAACCCGACGGGAATACCACTATTGGCA 157
1898 ThrGlnProAlaValLeuAlaGlyAlaArgValPhe.....LeuArgG1 1912
158 GCAGGGGGGAACCTTGGCAGCGCAGCGGTATCGGATTGGGAACATA 207
1912 nlySerSerThrSerLeuSer.AsnGly.....GlyAlaIle 1924
208 CAAGGCATCAGTTGGGCACACCTTTCATCCAGCAGCGCCATTAAAG 257
1925 GlnAsn.....GlnGlnSerGlyIleLysG1 1933
258 AAATATGGGGTAC..ATTGTCGCTTTTCGATCAGCGGACGAAATCC 304
1933 uSerValSerIrrpGlyThrGlnArgThrGlnAspAlaProHisGlnAla 1950
305 ATTCCCGCTTCGAC..AACCATGCTCAGATTCCGATTCTGTGAGACC 351
1950 rgProProLeuGlnArgGlyHisSerThrGlnIleProValGlyArgSer 1966
352 GGTAGTCCCGTTCGACGATTCAGCCTTACCAGTCATCCATTGGACGATA 401
1967 GlyAlaLeuAlaValAspValGlnMetGlnSerIleThrArgArgGlyPr 1983
402 CGAACACCATCCCGCGCAGCGCTATGACGGGCGCAGGGCGGCTATC 451
1983 oAspGlyGlnProGlnProGly..LeuGlnSerGlnGlyArgAlaAla... 1998
452 CGGCTCCCAAGCGCGAGGATATATACAGCTACGACATTAAGCGGTT 501
1999SerMe 2000
502 GCCCAAAATATCCGCTCAACCTGACGACAC..GCAGCACCG..G 545
2000 tPrArgLeuAlaIaGlnThrGlnProValThrAspAlaSerProMetL 2017
546 ACAAGCGCTTTCGACCGTTTCACATACCG.....GTAGTATGC 586
2017 ySerGlnSerIleSerThrLeuAlaGlnArgProArgGlyThrHisLeuGly 2033
587 TGACGCAAGAGTAGCGCAGGATTCAAACGCGCACCCGATACAGCCC 636
2034SerThrThrPr 2037
637 GAGCTGACAGATCGGCGCAATGCCCGGAAGCTTTCACGCGCACTGAGA 686
2037 O.....AspArgProProProSerGlnAlaSerSer.His..... 2048
687 TATGTCAAAAACATCATCGCGCGGAGGAGAAATGTCGCGCGCAGCG 726
2049HSHSHSHSHSArg.....CysHSHSHSHSArgatGat 2058
737 ATGCCGTGCAAGGTATTAAGCGAAAGCTCAACATTTGCTTATGACAGGC 786

2058 gasParGlys.....GlnArgSerLeuGlnLys..... 2067
787 TTGGGTCTGCTTTCACCCGAAACAGATGGCGGCATCAACGATT... 833
2068 ..GlyProSerLeuSerAlaAspMetAspGlyAlaProSerSerAlaVal 2083
834GGCAGATATGCGCGCACTCAAGACT 859
2084 GlyProGlyLeuProProGlyGlnGlyProThrGly..... 2095
860 ATCCGCGACAGCAGCATCCGCGATTGGGCAATCCAAACCCAGTCCGCA 909
2096 ..CysArgArgGlnArgGlnArg.....ArgGlnGlnArgS 2107
910 CAAGCATAGAACCGCTGACGAAATATCTTACGCGAGTCATCCCGTCAA 959
2107 eArgSerGlnGlnArgArgGlnProSerSerSerSerGlnLysGln 2123
960 AGGATTCGAGCTTTCGGGAAATACGCGCTTGGCGGC..... 999
2124 ArgPheTyrSerCysAsp.....Arg..PheGlyGlyArgGlnProP 2137
1000ATCAGGCAATCTGTGCAAGCGTTCGCAATG 1032
2137 rGlySerProLysProSerLeuSerSerHisProThrSerProThrAlaGly 2153
1033 GCGGAGATGCA..TTGCCGAAAGGAAATCCGCGTCACGCGACATTT 1079
2154 GlnGlnProGlyProHisProGlnLysSerGlySerValAsnGly.... 2168
1080 TCCGATGCGGCGATACGCCAATACCCGCTCCCTTACCATTTCCGCAATA 1129
2169SerProLeuLeuSerThrSer. 2175
1130 TCCGTTCAACTTGGAGCAGCGTTACGCGAAGAAACATCACTCTCA 1179
2176 ..GlyAlaSerThrProGlyArgGlyGlyArgArgGlnLeuProGlnThr 2191
1180 ACGGTCGCGCGTCAAAGGAAAGATGCAACCTGCAACACAAACAGCA 1229
2192 ProLeuThrProArgProSerIleThrTyrIleThrAlaAsn..... 2205
1230 CCCGAGAACCAAGTCCGTTGACGCT.....AAAGGTTTCCGAAT 1273
2206SerSerProIleHisPheAlaGlyAlaGlnThrSerLeuProAlaP 2221
1274 TTGAAAAAGAC...GTAAATACGATACGAGATTAATACGCTTACCA 1320
2221 heSerProGlyArgLeuSerArgGlyLeuSerGlnHisAsnAlaLeuLeu 2237
1321 CAAGTGAATCTATAGTAAACCGCTTAACTTAAGGTTCTGTGCG 1370
2238 GlnArgAspProLeuSerGlnProLeu..AlaProGlySerArgIleG1 2253
1371 ATCGGCTCAT.....TCTTGGCTTATACCTCCGAAATTCATTCGCA 1414
2253 ySerAspProGlyGlnGlyGlnArgLeuAspSerGlnAlaSerValHisA 2270
1415 AATTACA..... 1422
2270 lAlaLeuProGlnAspThrLeuThrPheGlnGlnAlaValAlaThrAsnSer 2286
1423 ...AGGCAAGTATGATGATGATATCCACCTAAATAATTTACTCTCTC 1469
2287 GlyArgSerSerArgThrSerTyrValSerSerLeuThr...SerGlnSe 2302
1470 AGCACCCTACCAAAAGCACTAATATGATATTTGGATTAATTTGGTA 1519
2302 rHisProLeuArgArgValPro..AsnGlyTyrHisCysThrLeuGly. 2317
1520 ATGAATGACATAAGCTCCATCAAGACTAAAGTCAAGAAATTTGAATGC 1569

